Learning new tricks:
Corticostriatal dynamics during novel skill learning

Fernando J. Santos

Dissertation presented to obtain the Ph.D degree in Biology | Neuroscience

Instituto de Tecnologia Química e Biológica António Xavier | Universidade Nova de Lisboa

Oeiras,
January, 2015

INSTITUTO
DE TECNOLOGIA
QUÍMICA E BIOLÓGICA
/UNL

Knowledge Creation

Learning new tricks:

Corticostriatal dynamics during novel skill learning

Fernando J. Santos

Dissertation presented to obtain the Ph.D degree in Biology | Neuroscience

Instituto de Tecnologia Química e Biológica | Universidade Nova de Lisboa

Research work coordinated by:

FUNDAÇÃO CALOUSTE GULBENKIAN
Instituto Gulbenkian de Ciência

Champalimaud Foundation

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA

Oeiras,
January, 2015

INSTITUTO
DE TECNOLOGIA
QUÍMICA E BIOLÓGICA
/UNL

Knowledge Creation

The work presented in this dissertation was carried out under the International Neuroscience Doctoral Programme (INDP, at the Champalimaud Neuroscience Programme, Champalimaud Centre for the Unknown, Lisbon, Portugal) under supervision of Dr. Rui M. Costa. Financial support was given by a doctoral fellowship from the Portuguese Science and Technology Foundation (FCT, Fundação para a Ciência e Tecnologia), with the reference SFRH/BD/33941/2009 attributed to Fernando João Antunes Gomes dos Santos.

ACKNOWLEDGEMENTS

To my parents, Fernando and Fernanda, who made all this possible by providing me the tools to thrive in any situation. Their example, their values and their education are undoubtedly the best skills I will ever acquire;

To my sisters, Daniela and Mariana, for their support and for always making me feel special even when I spent so much time far from home;

To Adriana, for all the patience and care during so many years of hard work and frustration;

To Cristiana, for bringing the best out of me, making me a much better and happier person, and for showing me that life can be so much fun;

To the Neurobiology of action folks, for all the discussions and insight, the long hours at the lab, but mostly for making me feel at home and surrounded by friends all the time;

To the CNP, an amazing group of people that managed to create a place and an environment where anyone would love to work;

To Rui, for being a mentor, an example and a friend.. and for making every piece of data, either good or bad, every experiment, either a sucessful or failed one, into something extraordinary, and more importantly, into a learning experience.

“If I have seen further it is by standing on the shoulders of giants”

Isaac Newton, in letter to Robert Hook (1676)

RESUMO

A aprendizagem de novas ações e habilidades motoras é uma capacidade prevalente em múltiplas espécies animais e uma característica crítica para a sobrevivência e competência num mundo em constante mudança. A geração e aprendizagem de novas ações ocorre através de um processo de tentativa e erro, no qual um animal explora o ambiente que o envolve, gerando múltiplos padrões de comportamento e selecionando aqueles que aumentam a probabilidade de consequências positivas. A adaptação e execução correta do padrão comportamental selecionado requer a coordenação de várias propriedades biomecânicas pelo animal, e os circuitos corticais e dos gânglios da base são conhecidos por estarem envolvidos nos processos relativos à aquisição, aprendizagem e consolidação de habilidades motoras.

Nestes estudos tentámos expandir o conhecimento relativo ao processo de aprendizagem e aquisição de uma habilidade motora nova. Utilizando uma tarefa operante difícil e complexa, demonstramos que os animais otimizam as componentes de uma sequência de ações que são relevantes para obter os resultados desejados. As componentes comportamentais são dinamicamente alteradas com base na sua relevância para a tarefa motora. Com o treino, há uma redução na variabilidade das componentes relevantes para a tarefa, que interferem diretamente com o objectivo da tarefa, não sendo esta redução observada ao nível da variabilidade das componentes não-relevantes, que não interferem com o objectivo da tarefa. A atividade dos circuitos córtex-estriado foi gravada de forma contínua, permitindo-nos seguir populações específicas de neurónios e avaliar alterações da sua

atividade que possam emergir com o treino de uma tarefa motora. Observámos que nas sessões iniciais há um aumento da variabilidade na atividade destes circuitos córtex-estriado, e que esta variabilidade diminui ao longo do processo de aprendizagem e das sessões de treino. Nestes estudos demonstrámos que estas dinâmicas neuronais estão especificamente correlacionadas com as dinâmicas das componentes comportamentais relevantes, que verificámos serem dependentes da existência de plasticidade sináptica nos circuitos córtex-estriado. A aprendizagem por tentativa e erro requer uma exploração inicial, de modo a analisar diferentes estratégias, resultando numa elevada variabilidade comportamental, putativamente causada por um aumento da variabilidade na atividade neuronal. O processo de aprendizagem promove uma adaptação gradual para uma estratégia de utilização de recursos, na qual o animal seleciona o padrão comportamental e neuronal que maximiza o resultado da tarefa. Nestes estudos testámos esta hipótese de seleção diretamente ao nível dos circuitos córtex-gânglios da base, utilizando um interface cérebro-máquina que nos permite reforçar um padrão de atividade neuronal específico. Verificámos que o emparelhamento explícito de um padrão neuronal experimentalmente definido, com uma ativação de neurónios dopaminérgicos, é suficiente para promover uma alteração na ocorrência deste padrão de atividade.

Em suma, os estudos apresentados nesta dissertação fornecem novas evidências para a teoria de seleção como base para a aprendizagem motora, integrando dados comportamentais e neuronais, de forma consistente com as teorias de feedback otimizado para controlo motor. Para além disso, a caracterização de uma tarefa motora de uma forma sistemática e menos dependente das especificações de cada paradigma, pode constituir uma estratégia útil para comparação de

resultados não só entre diferentes tarefas motoras, mas também entre tarefas de aprendizagem de memórias não-motoras.

SUMMARY

Learning novel actions and skills is a prevalent ability across multiple species and a critical feature for survival and competence in a constantly changing world. Novel actions are generated and learned through a process of trial and error, where an animal explores the environment around itself, generates multiple patterns of behavior and selects the ones that increase the likelihood of positive outcomes. Proper adaptation and execution of the selected behavior requires the coordination of several biomechanical features by the animal. Cortico-basal ganglia circuits and loops are critically involved in the acquisition, learning and consolidation of motor skills.

In these studies, we investigate the process of learning and acquisition of a novel motor skill. Using a difficult and complex operant task, we demonstrate that animals optimize the features of a sequence of actions that are relevant to obtain the desired outcome. Behavioral features are dynamically refined based on their relevance level to the task. Animals reduce variability in outcome-relevant features, which directly subserve the goals of the task, while there is no reduction in the variability of non-relevant features, which are unimportant for the task outcome. Activity of corticostriatal circuits was continuously recorded throughout learning, allowing us to follow specific populations of cells within these neuronal circuits and to monitor changes of activity that emerge with training of a motor skill. We demonstrate that in early sessions activity within corticostriatal circuits exhibit enhanced variability, and that this variability decreases as the animals progress through the training sessions. We show that neuronal dynamics are specifically correlated with the changes in the outcome-relevant behavioral features, which are

dependent on functional corticostriatal plasticity. These data provide support to the concepts of exploration and exploitation as basis for motor skill learning.

Trial and error learning requires initial exploration of different behavioral strategies, which results in increased behavioral variability, putatively caused by elevated neuronal variability. Learning promotes a gradual shift towards more exploitive strategies and a subsequent selection of behavioral and neuronal patterns that maximize the outcome of the task. We directly test this selection hypothesis within the cortico-basal ganglia networks by taking advantage of a brain-machine-interface paradigm to reinforce a specific spatiotemporal pattern. Pairing of an experimentally defined neuronal pattern with a phasic activation of dopaminergic neurons is sufficient to increase the likelihood of occurrence of this pattern.

Taken together, the studies presented in this dissertation provide further evidence of selection as a foundation for motor skill learning. The presented work integrates both behavioral and neuronal data in a framework consistent with optimal feedback as a general principle for motor control. Furthermore, characterization of an operant task in a systematic way, and less dependent on the specifics of each paradigm, might be a useful strategy to compare our results with different motor tasks, but also with other types of memory acquisition paradigms.

CONTENTS

ABBREVIATIONS LIST

BMI Brain-machine-interface

ChR2 Channelrhodopsin-2

CRF Continuous reinforcement

DLS Dorsolateral striatum

DMS Dorsomedial striatum

DS Dorsal striatum

FF Fano factor

FR Firing rate

GABA γ-aminobutyric acid

IPI Inter-press interval

ISI Inter-sequence interval

LMAN Lateral magnocellular nucleus of the anterior neostriatum

LTD Long-term depression

LTP Long-term potentiation

M1 Primary motor cortex

MSN Medium spiny neuron

NMDA N-methyl-D-aspartate

NSP Neural signal processor

PCA Principal components analysis

PETH Peri-event time histogram

SNc Substantia nigra pars compacta

VTA Ventral tegmental area

YFP Yellow fluorescent protein

1

INTRODUCTION

Congratulations! You have reached page 15! Now lets stop for a second to look backwards and realize how you have arrived here. Hopefully, browsing through these pages did not seem such a big challenge or achievement. But if you think about the complex patterns of movements that your fingers and hands had to perform, from grabbing and picking up this thesis from a table or shelf, to the perfect synchronization of finger muscles each time you flip a page, and the fact that all this was done with such ease and smoothness, are just small examples of the amazing capacity that we have to learn and acquire motor skills. We will go through the next pages, taking advantage of some of these motor skills that you have developed and mastered long ago, and attempt to understand how the process of learning a motor skill occurs, most of the time without us even being aware of it.

The ability to learn and perform motor skills is something that accompanies animals throughout life. From a small child babbling and discovering how to coordinate muscles and limbs in order to start walking or grabbing simple objects, to a more sophisticated stage where we learn how to ride bicycles and play the piano, the process of acquisition, selection and consolidation of novel motor skills, often time taken as granted by most of us, poses itself as an amazing challenge from the scientific point of view.

In 1953, Henry Molaison (widely known as patient H.M.) had parts of his

medial temporal lobe, hippocampus and amygdala removed in an attempt to cure his epilepsy, resulting in a case study that revolutionized the fields of neuroscience and our understanding of memory formation (Corkin, 2002). The case of H.M. is one of the first and most famous research studies that lead to the distinction between two types of memories: declarative (or explicit) and procedural (or implicit).

Declarative memory is the memory of facts and events that can be explicitly stored and retrieved, while procedural memory is the memory of skills and motor execution, such as using tools and moving our body, and is usually acquired through repetition and practice. After his surgery, which was successful in controlling the epilepsy, Molaison displayed a complete inability to form new long-term declarative memories, while maintaining a fully functional working-memory and ability to acquire novel motor skills and procedural memories (although without any explicit memory of having learned or practiced the task). This provided evidence that motor learning relies, at least partially, on brain structures that are separate from those involved in the acquisition of memories related to facts and events (Corkin, 2002).

Evidence from patient H.M. and several other amnesic patients have implicated brain regions like the hippocampus, entorhinal cortex, perirhinal and temporal cortices in the encoding and processing of declarative memories; whereas work in animal models and humans have implicated the cerebellum, basal ganglia and motor cortices of the frontal lobe in the acquisition and storage of procedural and motor memories (Doyon et al., 1997; Karni, 1996; Sanes and Donoghue, 2000) Evidence from patients with striatal dysfunctions due to Parkinson's or Huntington's disease, patients with cerebellar damage, and patients with lesions of the frontal motor cortices have given support to the role of cortico-striatal-thalamic and cortico-cerebellar-thalamic loops in motor processing (Doyon et al., 1997; Harrington et al., 2008; Heindel et al.,

1989).

Anatomical, neurophysiological and lesion studies in rodents and non-human primates have further demonstrated the importance and mechanisms through which these brain areas control motor performance and learning (Bolam et al., 2000; Middleton and Strick, 1997; 2000).

Motor learning: Behavior and neuronal dynamics

What exactly is motor learning? Although there is not a precise and consensual definition of motor learning, in most tasks and studies it is assessed by a reduction in reaction time and/or errors in motor performance, which is accomplished by changing movement synergies and kinematics. It is a process that occurs through repeated practice, by which movements get to be performed effortlessly and with improvements in accuracy, speed, and coordination (Guthrie, 1952; Schmidt and Lee, 2013; Welford, 1971; Willingham, 1998).

The concept of motor learning encompasses a wide range of learning paradigms: learning to control the gain of a reflex (Ito, 1993), improvement of a reaction time (Laubach et al., 2000), learning a finger tapping sequence (Nissen et al., 1987), or adjusting movements to external perturbations (Brashers-Krug et al., 1996). Skill learning is a specific type of motor learning, that relates to the acquisition of complex movements such as learning to play tennis or to ride a bicycle (Sanes, 2003). Learning complex movements is also usually designated as sequence learning, since the movements are often organized in sequences or chunks (Asanuma and Pavlides, 1997).

Most of the early research on motor learning focused on motor adaptation (Bedford, 1989; Bock, 1992; Caithness et al., 2004; Cunningham, 1989; Krakauer et al., 1999; 2000; Miall et al., 2004; Rabe

et al., 2009; Sainburg and Wang, 2002; Shadmehr and Mussa-Ivaldi, 1994; Shadmehr et al., 2010; Simani et al., 2007; Smith et al., 2006; Thoroughman and Shadmehr, 2000; Welch et al., 2007; Wolpert et al., 1995), which requires the subject to generate a modification in the motor output, with the goal of reducing systematic errors induced by external sensory or force perturbations. Adaptation could be achieved relatively rapidly using a forward model that makes adjustments of the motor commands based on sensory-prediction errors (Shadmehr et al., 2010).
Motor skill learning on the contrary, occurs in the absence of perturbation and its main goal is the reduction of variable performance errors (Deutsch and Newell, 2004; Guo and Raymond, 2010; Hung et al., 2008; Liu et al., 2006; Logan, 1988; Müller and Sternad, 2004; Ranganathan and Newell, 2010). Learning and performance are usually bound by the difficulty of the task and involve some form of speed-accuracy trade-off (Reis et al., 2009; Sanes et al., 1990).

Learning a motor skill can occur through a process of trial and error. By repetition, animals can explore their behavioral space and different strategies, evaluating the feedback from both correct and error trials, and selecting the pattern of movements that lead to the desired outcome.
There is growing evidence that skill learning occurs in different phases (Costa et al., 2004; Karni et al., 1998; Luft and Buitrago, 2005): an early phase, with fast improvements, facilitated by explicit knowledge (Stanley and Krakauer, 2013) where the memory is still labile and susceptible to competition and interference (Brashers-Krug et al., 1996; Walker et al., 2003); and a late phase, where the skill can be performed implicitly, evolving slowly and through repetition (Costa et al., 2004; Karni et al., 1998; Ungerleider et al., 2002), becoming consolidated and resistant to interference (Brashers-Krug et al., 1996; Walker et al., 2003).

Skill learning paradigms in both humans and animal models have shown improvements in motor performance, both during training sessions and during the intervals in between sessions (Buitrago et al., 2004a; 2004b; Karni et al., 1998), with a marked increase in performance when subjects are allowed to sleep (Walker et al., 2003). Mental rehearsal without practice can also lead to improvements in movements (Jeannerod, 1995; Mulder et al., 2004).
Several behavioral, electrophysiological, functional imaging and molecular experiments support these different stages of motor skill learning and their distinct behavioral and physiological hallmarks.

In 1973, Bliss and Lomo demonstrated for the first time the mechanism of long-term-potentiation (LTP), giving support for plasticity in neuronal circuits as one of the mechanisms for learning and memory (Bliss and Lomo, 1973). This mechanism has also been postulated to subserve motor learning, and since then, several studies have shown changes in activity and connectivity during motor learning in both motor cortex and striatum (Brasted and Wise, 2004; Classen et al., 1998; Costa et al., 2004; Debaere et al., 2004; Doyon and Benali, 2005; Gandolfo et al., 2000; Grafton et al., 1994; Jenkins et al., 1994; Kargo and Nitz, 2003; Karni et al., 1995; Kleim et al., 1998; 2002; Li et al., 2001; Muellbacher et al., 2002; Nudo et al., 1996; Seitz et al., 1990; Ungerleider et al., 2002; Wise et al., 1998; Yin et al., 2009). Striatal blockage of N-methyl-D-aspartate (NMDA) receptor mediated currents, which are critical for plasticity, severely impairs motor learning (Dang et al., 2006).

It has been widely observed that the activity of many motor cortical cells varies with muscle activity and kinematics (Cheney and Fetz, 1980; Evarts et al., 1983; Fromm, 1983; Humphrey et al., 1970; Thach, 1978), and that population activity of neuronal ensembles within the primary

motor cortex can code for movement features (Georgopoulos et al., 1986). In addition to the described dynamics during motor learning (Classen et al., 1998; Debaere et al., 2004; Gandolfo et al., 2000; Kargo and Nitz, 2003; Karni et al., 1995; Kleim et al., 1998; 2002; Li et al., 2001; Muellbacher et al., 2002; Nudo et al., 1996; Wise et al., 1998), differential activation of motor cortex throughout the different phases of motor learning has also been reported (Karni et al., 1998; Ungerleider et al., 2002). Large changes in topographic organization related to learning were observed within the primary motor cortex (M1) (Karni et al., 1995; Kleim et al., 1998; Nudo et al., 1996), and changes at the single cell level have also been documented (Wise et al., 1998). Network reorganizations in motor cortex are thought to occur through changes in synaptic efficacy induced by LTP (Iriki et al., 1989; Keller et al., 1990). Nevertheless, it is unreasonable to assume that motor cortex is the only site for plasticity in motor systems (Aizawa et al., 1991; Aosaki et al., 1994). Acquisition and execution of a skilled motor task requires the coordinated participation of a number of structures, including motor cortex, basal ganglia, cerebellum, and spinal cord.

The striatum, the major input nucleus of the basal ganglia, displays several changes in neural activity during motor and procedural learning (Barnes et al., 2005; Brasted and Wise, 2004; Carelli et al., 1997; Debaere et al., 2004; Doyon et al., 1996; Grafton et al., 1994; Jenkins et al., 1994; Jog et al., 1999; Seitz et al., 1990; Ungerleider et al., 2002), and the striatal circuits engaged during the early and late phases of skill learning also differ (Costa et al., 2004; Miyachi et al., 1997; 2002). As the entry point to the basal ganglia, the striatum stands in a relevant position to integrate information from the cortex, thalamus and modulations from the midbrain dopaminergic areas. The majority of neurons within the striatum (90-95%) (Kemp and Powell, 1971) are

medium spiny projections neurons (MSN), which are inhibitory GABA-ergic (γ-aminobutyric acid) cells (Kita and Kitai, 1988). MSNs receive excitatory, glutamatergic inputs mainly from the cortex but also from thalamus and amygdala (Voorn et al., 2004). The cortico-striatal-thalamic loops play an extremely important role in the control of goal-directed and habitual actions. Within the dorsal striatum of rodents, the most medial region (dorsomedial striatum, DMS, homologous to the caudate in primates) receives its major input from the associative cortical areas, while the most lateral region (dorsolateral striatum, DLS, homologous to the putamen in primates) receives its input from sensorimotor cortex (Haber, 2003; McGeorge and Faull, 1989; Voorn et al., 2004). The associative cortico-basal ganglia loop that courses through the DMS seems to be preferentially involved in the initial stages of visuomotor learning and during the rapid acquisition of action-outcome contingencies (Miyachi et al., 1997; 2002; Yin et al., 2005); while the sensorimotor cortico-basal ganglia loop that courses through the DLS is critical for the slower acquisition of automatic and habitual behaviors (Miyachi et al., 1997; 2002; Yin et al., 2004). Yin and colleagues (Yin et al., 2009) confirmed the differential involvement of the striatal areas in the different stages of skill learning through electrophysiology and lesion experiments, and have also observed the development of region- and pathway-specific plasticity. Within the striatum, ablation of NMDA receptors, which are essential for plasticity, was shown to disrupt the selection of behavioral patterns (Dang et al., 2006; Jin and Costa, 2010).

The cerebellum and the cortico-cerebellar-thalamic loops are also critical for some types of motor skill learning, with evidence of anatomical and functional differentiation between motor and associative regions, in similarity with the striatum (Middleton and Strick, 2000). While lesions of

the lateral cerebellar nuclei impair learning of sequences, visuomotor learning and spatial memory are not affected (Nixon and Passingham, 2000). On the other hand, blocking the dorsal part of the dentate nucleus, which projects to M1, does not impair learning new sequences, but disrupts the performance of learned sequences (Lu et al., 1998).

Basal ganglia and the cerebellum dependent learning is guided by reward and error signals (Doya, 2000). The striatum is densely innervated by midbrain dopaminergic inputs (Bolam et al., 2000), which are known to be critical for shaping the activity of the striatal circuits (Surmeier et al., 2011). Dopamine neurons display phasic reward / error signals that can be useful for learning and enhancing neuronal processes (Schultz, 2002). While the DMS receives projections from the ventral tegmental area (VTA) and ventromedial areas of the substantia nigra pars compacta (SNc), the DLS receives most of its dopaminergic projections from the dorsolateral SNc (Moore et al., 2001). Dopamine is known to be involved in both types of long-term synaptic plasticity: long-term potentiation (LTP) and long-term depression (LTD) (Gerfen and Surmeier, 2011; Shen et al., 2008). On one hand, plasticity between the cortical projection neurons and the MSNs seems to occur more easily at the DMS in the form of LTP, by a process dependent on the activation of dopamine type-1 (D1) receptors and NMDA glutamate receptors (Kerr and Wickens, 2001; Partridge et al., 2000); on the other hand, LTD was found to be more easily induced in the DLS, by a process dependent on dopaminergic and endocannabinoid signaling (Gerdeman et al., 2002; Kreitzer and Malenka, 2005; Partridge et al., 2000). During motor skill learning there is also differential plasticity within the two major output pathways of the striatum, with an increase of LTP in dopamine type-2 (D2) receptor-expressing cells, which are mainly the MSNs projecting to the substantia nigra through the globus pallidus (striatopallidal pathway),

and a decrease in the dependency of activation of D1 receptors, which are mainly expressed in the pathway projecting directly to the substantia nigra (striatonigral pathway) (Yin et al., 2009). Furthermore, high dopamine levels, which are associated with hyperkinesia and increased exploration, lead to less correlation in neural activity and more asynchronous activity in cortico-basal ganglia network (Costa et al., 2006); on the contrary, depletion of dopamine, which is known to disrupt skillful performance of sequential movements (Matsumoto et al., 1999) and reduce the occurrence of voluntary movements, leads to an increase of correlated activity and synchronous activity within these networks (Costa et al., 2006). Work in songbirds has suggested that dopamine is implicated in the modulation of variability both at the level of neuronal activity and motor output (Gale and Perkel, 2010).

While in the basal ganglia, cortical inputs are combined at the level of MSNs with signals carried by dopaminergic neurons; in the cerebellum, Purkinje cells integrate indirect cortical inputs with sensorimotor error signals, mediated by the climbing fibers (Wang et al., 2000). These signals are then transmitted back to the cortex through these corticostriatal and cortico-cerebellar loops, which creates a feedback process that is likely critical for motor skill learning.

Motor and neuronal variability during motor learning

As described above, several studies support the idea that acquisition of skilled movements happens through a process of trial and error. When animals are learning a novel motor task, repetition of the desired movement is never exactly the same and has some intrinsic variability associated to it. These fluctuations in motor performance are usually thought as inevitable and undesirable, and can have a neuronal, neuromuscular, musculoskeletal or even environmental origin

(Churchland et al., 2006; Jones et al., 2002; Osborne et al., 2005; Schmidt et al., 1979; Stein et al., 2005). Throughout the brain, there is a high degree of variability in the activity of neurons, even in constant task conditions (Arieli et al., 1996; Bach and Krüger, 1986; Fox and Raichle, 2007; Lee et al., 1998; Plenz, 2012; Vogels et al., 1989). There are multiple contributions to this neuronal variability: from the inherent noise of cellular, molecular and synaptic mechanisms, to the noise introduced by uncertain sensory inputs and erratic muscle activity (Arieli et al., 1996; Bialek and Setayeshgar, 2008; Faisal et al., 2008; Stein et al., 2005; Vogels et al., 1989).

Therefore, variability in neuronal activity and in behavior output is usually treated as noise and a cause for poor performance. Motor variability is in general considered as signal-dependent noise, varying proportionally to the magnitude of the motor output (Jones et al., 2002). This view treats variability as detrimental for the neuronal computations, and supports the hypothesis that the objective of the brain during learning would be to minimize variability to generate a better readout of the information coded within the neuronal ensembles (Bialek et al., 1991; Shadlen and Newsome, 1998). Several theories of motor control postulate that actions are organized and selected specifically to minimize the extent to which variability affects performance (Harris and Wolpert, 1998; O'Sullivan et al., 2009; Scholz and Schöner, 1999; Todorov, 2004; Todorov and Jordan, 2002; van Beers et al., 2002). In agreement with this hypothesis, studies have shown high levels of variability both at a behavioral (Jin and Costa, 2010; Miller et al., 2010; Tumer and Brainard, 2007) and neuronal level (Barnes et al., 2005; Costa et al., 2004) in early learning phases, with subsequent reduction of neuronal variability within the corticostriatal circuits, as motor memories are consolidated and behavior becomes less variable (Barnes et al., 2005; Costa et al.,

2004; Jin and Costa, 2010; Kao et al., 2005).

However, recent studies have put forth the hypothesis that variability might not purely reflect noise, but rather might serve as a part of the neuronal signal. Indeed, the brain can generate and increase variability as a critical feature of some learning paradigms (Faisal et al., 2008; Maimon and Assad, 2009; Stein et al., 2005). This variability might arise from inherently variable characteristics of the motor system, but could nevertheless subserve a critical function: the exploration of the plethora of behavioral possibilities that an organism can generate (Costa, 2011; Friston, 2010). Variability can be important not only for motor learning (Faisal et al., 2008; Fiete et al., 2007; Rokni et al., 2007) but also for more general learning mechanisms (Fusi, 2002), by allowing for the exploration of different motor programs and neuronal patterns that lead to the desired output.

Research on bird song learning supports this hypothesis of variability as a critical player in motor learning (Fee and Goldberg, 2011). The lateral magnocellular nucleus of the anterior neostriatum (LMAN) is a nucleus that is part of the anterior forebrain pathway, which is a basal ganglia-dorsal forebrain circuit with an important role in vocal learning. The LMAN is critical for motor performance and learning ability of songbirds (Charlesworth et al., 2012; Kao et al., 2005; Olveczky et al., 2005). The LMAN projects to a cortical output area involved in singing and is responsible for generating the variability necessary to promote learning (Olveczky et al., 2005; Tumer and Brainard, 2007).
Neuronal variability modulations are also observed during motor planning, such that they increased immediately before movement onset even without any changes in external noise (Churchland et al., 2006). The generation of higher motor variability is hypothesized to create

higher behavioral exploration, which is an important feature of reinforcement learning and to allow an animal to better probe the environment and the possibilities regarding the task at hand (Kaelbling et al., 1996; Sutton and Barto, 1998). Motor variability can therefore facilitate motor learning and the nervous system can actively modulate it in order to improve learning (Wu et al., 2014).

The reduction and minimization of variability could be seen as a general functional mechanism for sensory and motor areas (Harris and Wolpert, 1998; Todorov, 2005; van Beers et al., 2004). Despite the considerable trial-to-trial variability, some features of motor performance are tightly controlled, particularly features that are relevant for the task (Scholz and Schöner, 1999). The motor system can thus learn complex tasks by optimizing variability in the dimensions that are relevant for the task (Diedrichsen et al., 2010; Scott, 2004; Todorov and Jordan, 2002; Valero-Cuevas et al., 2009) and allowing variability to accumulate in dimensions that are not relevant for the task at hand, a process consistent with “optimal feedback” as a theory for motor control (Todorov and Jordan, 2002). Todorov and Jordan (Todorov and Jordan, 2002) named this the “minimal intervention principle”, arguing that deviations from the desired movement are corrected in a selective way, only when they interfere with task performance. This model also predicts that in situations were noise is perturbing the motor system homogenously, variability will be higher in task-irrelevant dimensions, which will not be corrected by motor control (Todorov and Jordan, 2002). Motor variability is critical for these motor adjustments (Tumer and Brainard, 2007), with continouous accurate feedback being essencial for correct adaptation (Sakata and Brainard, 2006).

Brain-machine-interfaces as paradigms for understanding the neuronal dynamics of learning

Orsborn and Carmena (Orsborn and Carmena, 2013) have recently reviewed evidence to demonstrate how brain-machine-interface (BMI) paradigms with closed-loop designs can be used to better understand neural changes underlying skill learning. In BMI tasks recordings of real time neuronal activity are used to control some external variable. Feedback regarding the control signal is provided to the subject, either by external sensory stimulation or direct brain activation, allowing him to control the neuronal activity that is used as input for the BMI controller.

Several studies have shown that the firing patterns of individual cells (Fetz, 1969) and ensembles of cells (Clancy et al., 2014; Koralek et al., 2012) can be conditioned and volitionally controlled by animals, with the existing structure within the neuronal networks able to facilitate learning (Sadtler et al., 2014). Regardless of differences between BMI and natural sensorimotor learning, BMI closed-loop systems can still promote learning and adaptation (Fetz, 2007) and rely on similar components of the nervous system.

Therefore, BMIs are particularly useful for interrogations of the neurophysiological mechanisms of motor and skill learning, due to the inherent characteristic of being controlled by the experimenter and allowing a reduction in the complexity associated to natural learning.

As suggested by the aforementioned studies, the natural learning process requires exploration of a behavioral repertoire and their respective outcomes, with reinforcement of motor actions that lead to preferred outcomes (Sutton and Barto, 1998). These observations fully support Thorndike's law of effect, which stated that behavioral responses that were most closely followed by a satisfying result were most likely to become established patterns and to occur again (Thorndike, 1898).

Acquiring and learning novel skills implies that networks of neurons are able to generate new activity patterns more often and reliably. If the contingency between a specific behavioral pattern and a reinforcement is enough to bias the selection of that specific behavior and increase its likelihood of occurrence, then in theory, pairing a specific pattern of neuronal activity with reinforcement might lead to selection of that same pattern and concurrent increase in its prevalence.

AIMS: Neuronal and behavior dynamics of action learning

It is recognized the importance of variability both from a behavioral and neuronal point of view for motor control. Nevertheless its impact onto the learning process is still poorly understood. Aditional knowledge regarding the mechanisms and dynamics of motor skill learning from a cortical, subcortical and behavioral level could help us to better tackle some of the challenges regarding learning and performance of motor skills. Therefore in the next chapters we aim to:

- Investigate the process of learning a novel and complex motor skill, from the behavioral to the neuronal level, evaluating the dynamics of behavioral and neuronal variability throughout the different learning stages;

- Evaluate if the direct reinforcement of a specific neuronal activity pattern, using a brain-machine-interface closed-loop paradigm to stimulate midbrain dopaminergic neurons, is sufficient to promote the selection of that neuronal pattern, similarly to the dynamics observed during natural motor skill learning.

REFERENCES

Aizawa, H., Inase, M., Mushiake, H., Shima, K., and Tanji, J. (1991). Reorganization of activity in the supplementary motor area associated with motor learning and functional recovery. Exp. Brain Res. *84*, 668–671.

Aosaki, T., Tsubokawa, H., Ishida, A., Watanabe, K., Graybiel, A.M., and Kimura, M. (1994). Responses of tonically active neurons in the primate's striatum undergo systematic changes during behavioral sensorimotor conditioning. J. Neurosci. *14*, 3969–3984.

Arieli, A., Sterkin, A., Grinvald, A., and Aertsen, A. (1996). Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. Science *273*, 1868–1871.

Asanuma, H., and Pavlides, C. (1997). Neurobiological basis of motor learning in mammals. Neuroreport *8*, i–vi.

Bach, M., and Krüger, J. (1986). Correlated neuronal variability in monkey visual cortex revealed by a multi-microelectrode. Exp. Brain Res. *61*, 451–456.

Barnes, T.D., Kubota, Y., Hu, D., Jin, D.Z., and Graybiel, A.M. (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. Nature *437*, 1158–1161.

Bedford, F.L. (1989). Constraints on learning new mappings between perceptual dimensions. J Exp Psychol Hum Percept Perform *15*, 232–248.

Bialek, W., Rieke, F., de Ruyter van Steveninck, R.R., and Warland, D. (1991). Reading a neural code. Science *252*, 1854–1857.

Bialek, W., and Setayeshgar, S. (2008). Cooperativity, sensitivity, and noise in biochemical signaling. Phys. Rev. Lett. *100*, 258101.

Bliss, T.V., and Lomo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. J Physiol (Lond) *232*, 331–356.

Bock, O. (1992). Adaptation of aimed arm movements to sensorimotor discordance: evidence for direction-independent gain control. Behavioural Brain Research *51*, 41–50.

Bolam, J.P., Hanley, J.J., Booth, P.A., and Bevan, M.D. (2000). Synaptic organisation of the basal ganglia. J. Anat. *196 (Pt 4)*, 527–542.

Brashers-Krug, T., Shadmehr, R., and Bizzi, E. (1996). Consolidation in human motor memory. Nature *382*, 252–255.

Brasted, P.J., and Wise, S.P. (2004). Comparison of learning-related neuronal activity in the dorsal premotor cortex and striatum. Eur J Neurosci *19*, 721–740.

Buitrago, M.M., Ringer, T., Schulz, J.B., Dichgans, J., and Luft, A.R. (2004a). Characterization of motor skill and instrumental learning time scales in a skilled reaching task in rat. Behavioural Brain Research *155*, 249–256.

Buitrago, M.M., Schulz, J.B., Dichgans, J., and Luft, A.R. (2004b). Short and long-term motor skill learning in an accelerated rotarod training paradigm. Neurobiol Learn Mem *81*, 211–216.

Caithness, G., Osu, R., Bays, P., Chase, H., Klassen, J., Kawato, M., Wolpert, D.M., and Flanagan, J.R. (2004). Failure to consolidate the consolidation theory of learning for sensorimotor adaptation tasks. J. Neurosci. *24*, 8662–8671.

Carelli, R.M., Wolske, M., and West, M.O. (1997). Loss of lever press-related firing of rat striatal forelimb neurons after repeated sessions in a lever pressing task. J. Neurosci. *17*, 1804–1814.

Charlesworth, J.D., Warren, T.L., and Brainard, M.S. (2012). Covert skill learning in a cortical-basal ganglia circuit. Nature *486*, 251–255.

Cheney, P.D., and Fetz, E.E. (1980). Functional classes of primate corticomotoneuronal cells and their relation to active force. J. Neurophysiol. *44*, 773–791.

Churchland, M.M., Afshar, A., and Shenoy, K.V. (2006). A Central Source of Movement Variability. Neuron *52*, 1085–1096.

Clancy, K.B., Koralek, A.C., Costa, R.M., Feldman, D.E., and Carmena, J.M. (2014). Volitional modulation of optically recorded calcium signals during neuroprosthetic learning. Nat. Neurosci. *17*, 807–809.

Classen, J., Liepert, J., Wise, S.P., Hallett, M., and Cohen, L.G. (1998). Rapid plasticity of human cortical movement representation induced by practice. J. Neurophysiol. *79*, 1117–1123.

Corkin, S. (2002). What's new with the amnesic patient H.M.? Nat. Rev. Neurosci. *3*, 153–160.

Costa, R.M. (2011). A selectionist account of de novo action learning. Curr. Opin. Neurobiol. *21*, 579–586.

Costa, R.M., Cohen, D., and Nicolelis, M.A.L. (2004). Differential corticostriatal plasticity during fast and slow motor skill learning in mice. Curr. Biol. *14*, 1124–1134.

Costa, R.M., Lin, S.-C., Sotnikova, T.D., Cyr, M., Gainetdinov, R.R., Caron, M.G., and Nicolelis, M.A.L. (2006). Rapid alterations in corticostriatal ensemble coordination during acute dopamine-dependent motor dysfunction. Neuron *52*, 359–369.

Cunningham, H.A. (1989). Aiming error under transformed spatial mappings suggests a structure for visual-motor maps. J Exp Psychol Hum Percept Perform *15*, 493–506.

Dang, M.T., Yokoi, F., Yin, H.H., Lovinger, D.M., Wang, Y., and Li, Y. (2006). Disrupted motor learning and long-term synaptic plasticity in mice lacking NMDAR1 in the striatum. Proc. Natl. Acad. Sci. USA *103*, 15254–15259.

Debaere, F., Wenderoth, N., Sunaert, S., Van Hecke, P., and Swinnen, S.P. (2004). Changes in brain activation during the acquisition of a new bimanual coodination task. Neuropsychologia *42*, 855–867.

Deutsch, K.M., and Newell, K.M. (2004). Changes in the structure of children's isometric force variability with practice. J Exp Child Psychol *88*, 319–333.

Diedrichsen, J., Shadmehr, R., and Ivry, R.B. (2010). The coordination of movement: optimal feedback control and beyond. Trends Cogn. Sci. *14*, 31–39.

Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. Curr. Opin. Neurobiol. *10*, 732–739.

Doyon, J., and Benali, H. (2005). Reorganization and plasticity in the adult brain during learning of motor skills. Curr. Opin. Neurobiol. *15*, 161–167.

Doyon, J., Gaudreau, D., Laforce, R., Castonguay, M., Bédard, P.J., Bédard, F., and Bouchard, J.P. (1997). Role of the striatum, cerebellum, and frontal lobes in the learning of a visuomotor sequence. Brain and Cognition *34*, 218–245.

Doyon, J., Owen, A.M., Petrides, M., Sziklas, V., and Evans, A.C. (1996). Functional anatomy of visuomotor skill learning in human subjects examined with positron emission tomography. *8*, 637–648.

Evarts, E.V., Fromm, C., Kroller, J., and Jennings, V.A. (1983). Motor Cortex control of finely graded forces. J. Neurophysiol. *49*, 1199–1215.

Faisal, A.A., Selen, L.P.J., and Wolpert, D.M. (2008). Noise in the nervous system. Nat. Rev. Neurosci. *9*, 292–303.

Fee, M.S., and Goldberg, J.H. (2011). A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. Neurosci. *198*, 152–170.

Fetz, E.E. (1969). Operant conditioning of cortical unit activity. Science *163*, 955–958.

Fetz, E.E. (2007). Volitional control of neural activity: implications for brain-computer interfaces. J Physiol (Lond) *579*, 571–579.

Fiete, I.R., Fee, M.S., and Seung, H.S. (2007). Model of birdsong learning based on gradient

estimation by dynamic perturbation of neural conductances. J. Neurophysiol. *98*, 2038–2057.

Fox, M.D., and Raichle, M.E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. Nat. Rev. Neurosci. *8*, 700–711.

Friston, K. (2010). The free-energy principle: a unified brain theory? Nat. Rev. Neurosci. *11*, 127–138.

Fromm, C. (1983). Changes of steady state activity in motor cortex consistent with the length-tension relation of muscle. Pflugers Arch. *398*, 318–323.

Fusi, S. (2002). Hebbian spike-driven synaptic plasticity for learning patterns of mean firing rates. Biol Cybern *87*, 459–470.

Gale, S.D., and Perkel, D.J. (2010). A basal ganglia pathway drives selective auditory responses in songbird dopaminergic neurons via disinhibition. J. Neurosci. *30*, 1027–1037.

Gandolfo, F., Li, C., Benda, B.J., Schioppa, C.P., and Bizzi, E. (2000). Cortical correlates of learning in monkeys adapting to a new dynamical environment. Proc. Natl. Acad. Sci. USA *97*, 2259–2263.

Georgopoulos, A.P., Schwartz, A.B., and Kettner, R.E. (1986). Neuronal population coding of movement direction. Science *233*, 1416–1419.

Gerdeman, G.L., Ronesi, J., and Lovinger, D.M. (2002). Postsynaptic endocannabinoid release is critical to long-term depression in the striatum. Nat. Neurosci. *5*, 446–451.

Gerfen, C.R., and Surmeier, D.J. (2011). Modulation of striatal projection systems by dopamine. Annu. Rev. Neurosci. *34*, 441–466.

Grafton, S.T., Woods, R.P., and Tyszka, M. (1994). Functional imaging of procedural motor learning: Relating cerebral blood flow with individual subject performance. Hum Brain Mapp *1*, 221–234.

Guo, C.C., and Raymond, J.L. (2010). Motor learning reduces eye movement variability through reweighting of sensory inputs. J. Neurosci. *30*, 16241–16248.

Guthrie, E.R. (1952). The psychology of learning (Rev. ed.). (Oxford: Harper).

Haber, S.N. (2003). The primate basal ganglia: parallel and integrative networks. Journal of Chemical Neuroanatomy *26*, 317–330.

Harrington, D.L., Haaland, K.Y., Yeo, R.A., and Marder, E. (2008). Procedural memory in Parkinson's disease: Impaired motor but not visuoperceptual learning. Journal of Clinical and Experimental Neuropsychology *12*, 323–339.

Harris, C.M., and Wolpert, D.M. (1998). Signal-dependent noise determines motor planning. Nature *394*, 780–784.

Heindel, W.C., Salmon, D.P., Shults, C.W., Walicke, P.A., and Butters, N. (1989). Neuropsychological evidence for multiple implicit memory systems: A comparison of Alzheimer's, Huntington‘s, and Parkinson’s disease patients. J. Neurosci. *9*, 582–587.

Humphrey, D.R., Schmidt, E.M., and Thompson, W.D. (1970). Predicting measures of motor performance from multiple cortical spike trains. Science *170*, 758–762.

Hung, Y.-C., Kaminski, T.R., Fineman, J., Monroe, J., and Gentile, A.M. (2008). Learning a multi-joint throwing task: a morphometric analysis of skill development. Exp. Brain Res. *191*, 197–208.

Iriki, A., Pavlides, C., Keller, A., and Asanuma, H. (1989). Long-term potentiation in the motor cortex. Science *245*, 1385–1387.

Ito, M. (1993). Neurophysiology of the noduloflocculars system. Rev. Neurol. (Paris) *149*, 692–697.

Jeannerod, M. (1995). Mental imagery in the motor context. Neuropsychologia *33*, 1419–1432.

Jenkins, I.H., Brooks, D.J., Nixon, P.D., Frackowiak, R.S., and Passingham, R.E. (1994). Motor sequence learning: a study with positron emission tomography. J. Neurosci. *14*, 3775–3790.

Jin, X., and Costa, R.M. (2010). Start/stop signals emerge in nigrostriatal circuits during sequence

learning. Nature *466*, 457–462.

Jog, M.S., Kubota, Y., Connolly, C.I., Hillegaart, V., and Graybiel, A.M. (1999). Building neural representations of habits. Science *286*, 1745–1749.

Jones, K.E., Hamilton, A.F., and Wolpert, D.M. (2002). Sources of signal-dependent noise during isometric force production. J. Neurophysiol. *88*, 1533–1544.

Kaelbling, L.P., Littman, M.L., and Moore, A.W. (1996). Reinforcement Learning: A Survey. Journal of Artificial Intelligence Research 237–285.

Kao, M.H., Doupe, A.J., and Brainard, M.S. (2005). Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. Nature *433*, 638–643.

Kargo, W.J., and Nitz, D.A. (2003). Early skill learning is expressed through selection and tuning of cortically represented muscle synergies. J. Neurosci. *23*, 11255–11269.

Karni, A. (1996). The acquisition of perceptual and motor skills: a memory system in the adult human cortex. Brain Res Cogn Brain Res *5*, 39–48.

Karni, A., Meyer, G., Jezzard, P., Adams, M.M., Turner, R., and Ungerleider, L.G. (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. Nature *377*, 155–158.

Karni, A., Meyer, G., Rey-Hipolito, C., Jezzard, P., Adams, M.M., Turner, R., and Ungerleider, L.G. (1998). The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex. Proc. Natl. Acad. Sci. USA *95*, 861–868.

Keller, A., Iriki, A., and Asanuma, H. (1990). Identification of neurons producing long-term potentiation in the cat motor cortex: intracellular recordings and labeling. J. Comp. Neurol. *300*, 47–60.

Kemp, J.M., and Powell, T.P. (1971). The connexions of the striatum and globus pallidus: synthesis and speculation. Philos. Trans. R. Soc. Lond., B, Biol. Sci. *262*, 441–457.

Kerr, J.N., and Wickens, J.R. (2001). Dopamine D-1/D-5 receptor activation is required for long-term potentiation in the rat neostriatum in vitro. J. Neurophysiol. *85*, 117–124.

Kita, H., and Kitai, S.T. (1988). Glutamate decarboxylase immunoreactive neurons in rat neostriatum: their morphological types and populations. Brain Res. *447*, 346–352.

Kleim, J.A., Barbay, S., and Nudo, R.J. (1998). Functional reorganization of the rat motor cortex following motor skill learning. J. Neurophysiol. *80*, 3321–3325.

Kleim, J.A., Barbay, S., Cooper, N.R., Hogg, T.M., Reidel, C.N., Remple, M.S., and Nudo, R.J. (2002). Motor learning-dependent synaptogenesis is localized to functionally reorganized motor cortex. Neurobiol Learn Mem *77*, 63–77.

Koralek, A.C., Jin, X., Long, J.D., Costa, R.M., and Carmena, J.M. (2012). Corticostriatal plasticity is necessary for learning intentional neuroprosthetic skills. Nature *483*, 331–335.

Krakauer, J.W., Ghilardi, M.F., and Ghez, C. (1999). Independent learning of internal models for kinematic and dynamic control of reaching. Nat. Neurosci. *2*, 1026–1031.

Krakauer, J.W., Pine, Z.M., Ghilardi, M.F., and Ghez, C. (2000). Learning of visuomotor transformations for vectorial planning of reaching trajectories. J. Neurosci. *20*, 8916–8924.

Kreitzer, A.C., and Malenka, R.C. (2005). Dopamine modulation of state-dependent endocannabinoid release and long-term depression in the striatum. J. Neurosci. *25*, 10537–10545.

Laubach, M., Wessberg, J., and Nicolelis, M.A.L. (2000). Cortical ensemble activity increasingly predicts behaviour outcomes during learning of a motor task. Nature *405*, 567–571.

Lee, D., Port, N.L., Kruse, W., and Georgopoulos, A.P. (1998). Variability and correlated noise in the discharge of neurons in motor and parietal areas of the primate cortex. J. Neurosci. *18*, 1161–1170.

Li, C.-S.R., Padoa-Schioppa, C., and Bizzi, E. (2001). Neuronal correlates of motor performance

and motor learning in the primary motor cortex of monkeys adapting to an external force field. Neuron *30*, 593–607.

Liu, Y.-T., Mayer-Kress, G., and Newell, K.M. (2006). Qualitative and quantitative change in the dynamics of motor learning. J Exp Psychol Hum Percept Perform *32*, 380–393.

Logan, G.D. (1988). Toward an instance theory of automatization. Psychol Rev *95*, 492–527.

Lu, X., Hikosaka, O., and Miyachi, S. (1998). Role of monkey cerebellar nuclei in skill for sequential movement. J. Neurophysiol. *79*, 2245–2254.

Luft, A.R., and Buitrago, M.M. (2005). Stages of motor skill learning. Mol Neurobiol *32*, 205–216.

Maimon, G., and Assad, J.A. (2009). Beyond Poisson: increased spike-time regularity across primate parietal cortex. Neuron *62*, 426–440.

Matsumoto, N., Hanakawa, T., Maki, S., Graybiel, A.M., and Kimura, M. (1999). Role of nigrostriatal dopamine system in learning to perform sequential motor tasks in a predictive manner. J. Neurophysiol. *82*, 978–998.

McGeorge, A.J., and Faull, R.L. (1989). The organization of the projection from the cerebral cortex to the striatum in the rat. Neurosci. *29*, 503–537.

Miall, R.C., Jenkinson, N., and Kulkarni, K. (2004). Adaptation to rotated visual feedback: a re-examination of motor interference. Exp. Brain Res. *154*, 201–210.

Middleton, F.A., and Strick, P.L. (1997). Cerebellar output channels. Int. Rev. Neurobiol. *41*, 61–82.

Middleton, F.A., and Strick, P.L. (2000). Basal ganglia and cerebellar loops: motor and cognitive circuits. Brain Res. Brain Res. Rev. *31*, 236–250.

Miller, J.E., Hilliard, A.T., and White, S.A. (2010). Song practice promotes acute vocal variability at a key stage of sensorimotor learning. PLoS ONE *5*, e8592.

Miyachi, S., Hikosaka, O., Miyashita, K., Kárádi, Z., and Rand, M.K. (1997). Differential roles of monkey striatum in learning of sequential hand movement. Exp. Brain Res. *115*, 1–5.

Miyachi, S., Hikosaka, O., and Lu, X. (2002). Differential activation of monkey striatal neurons in the early and late stages of procedural learning. Exp. Brain Res. *146*, 122–126.

Moore, A.E., Cicchetti, F., Hennen, J., and Isacson, O. (2001). Parkinsonian motor deficits are reflected by proportional A9/A10 dopamine neuron degeneration in the rat. Experimental Neurology *172*, 363–376.

Muellbacher, W., Ziemann, U., Wissel, J., Dang, N., Kofler, M., Facchini, S., Boroojerdi, B., Poewe, W., and Hallett, M. (2002). Early consolidation in human primary motor cortex. Nature *415*, 640–644.

Mulder, T., Zijlstra, S., Zijlstra, W., and Hochstenbach, J. (2004). The role of motor imagery in learning a totally novel movement. Exp. Brain Res. *154*, 211–217.

Müller, H., and Sternad, D. (2004). Decomposition of variability in the execution of goal-oriented tasks: three components of skill improvement. J Exp Psychol Hum Percept Perform *30*, 212–233.

Nissen, M.J., Knopman, D.S., and Schacter, D.L. (1987). Neurochemical dissociation of memory systems. Neurology *37*, 789–794.

Nixon, P.D., and Passingham, R.E. (2000). The cerebellum and cognition: cerebellar lesions impair sequence learning but not conditional visuomotor learning in monkeys. Neuropsychologia *38*, 1054–1072.

Nudo, R.J., Milliken, G.W., Jenkins, W.M., and Merzenich, M.M. (1996). Use-dependent alterations of movement representations in primary motor cortex of adult squirrel monkeys. J. Neurosci. *16*, 785–807.

O'Sullivan, I., Burdet, E., and Diedrichsen, J. (2009). Dissociating variability and effort as determinants of coordination. PLoS Comp Biol *5*, e1000345.

Olveczky, B.P., Andalman, A.S., and Fee, M.S. (2005). Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. PLoS Biol. *3*, e153.

Orsborn, A.L., and Carmena, J.M. (2013). Creating new functional circuits for action via brain-machine interfaces. Front Comput Neurosci *7*, 157.

Osborne, L.C., Lisberger, S.G., and Bialek, W. (2005). A sensory source for motor variation. Nature *437*, 412–416.

Partridge, J.G., Tang, K.C., and Lovinger, D.M. (2000). Regional and postnatal heterogeneity of activity-dependent long-term changes in synaptic efficacy in the dorsal striatum. J. Neurophysiol. *84*, 1422–1429.

Plenz, D. (2012). Neuronal avalanches and coherence potentials. Eur. Phys. J. Spec. Top. *205*, 259–301.

Rabe, K., Livne, O., Gizewski, E.R., Aurich, V., Beck, A., Timmann, D., and Donchin, O. (2009). Adaptation to visuomotor rotation and force field perturbation is correlated to different brain areas in patients with cerebellar degeneration. J. Neurophysiol. *101*, 1961–1971.

Ranganathan, R., and Newell, K.M. (2010). Influence of motor learning on utilizing path redundancy. Neurosci. Lett. *469*, 416–420.

Reis, J., Schambra, H.M., Cohen, L.G., Buch, E.R., Fritsch, B., Zarahn, E., Celnik, P.A., and Krakauer, J.W. (2009). Noninvasive cortical stimulation enhances motor skill acquisition over multiple days through an effect on consolidation. Proc. Natl. Acad. Sci. USA *106*, 1590–1595.

Rokni, U., Richardson, A.G., Bizzi, E., and Seung, H.S. (2007). Motor learning with unstable neural representations. Neuron *54*, 653–666.

Sadtler, P.T., Quick, K.M., Golub, M.D., Chase, S.M., Ryu, S.I., Tyler-Kabara, E.C., Yu, B.M., and Batista, A.P. (2014). Neural constraints on learning. Nature *512*, 423–426.

Sainburg, R.L., and Wang, J. (2002). Interlimb transfer of visuomotor rotations: independence of direction and final position information. Exp. Brain Res. *145*, 437–447.

Sakata, J.T., and Brainard, M.S. (2006). Real-time contributions of auditory feedback to avian vocal motor control. J. Neurosci. *26*, 9619–9628.

Sanes, J.N., and Donoghue, J.P. (2000). Plasticity and primary motor cortex. Annu. Rev. Neurosci. *23*, 393–415.

Sanes, J.N., Dimitrov, B., and Hallett, M. (1990). Motor learning in patients with cerebellar dysfunction. Brain *113 (Pt 1)*, 103–120.

Sanes, J.N. (2003). Neocortical mechanisms in motor learning. Curr. Opin. Neurobiol. *13*, 225–231.

Schmidt, R.A., Zelaznik, H., Hawkins, B., Frank, J.S., and Quinn, J.T. (1979). Motor-output variability: a theory for the accuracy of rapid motor acts. Psychol Rev *47*, 415–451.

Schmidt, R., and Lee, T. (2013). Motor Learning and Performance, 5E With Web Study Guide (Human Kinetics).

Scholz, J.P., and Schöner, G. (1999). The uncontrolled manifold concept: identifying control variables for a functional task. Exp. Brain Res. *126*, 289–306.

Schultz, W. (2002). Getting formal with dopamine and reward. Neuron *36*, 241–263.

Scott, S.H. (2004). Optimal feedback control and the neural basis of volitional motor control. Nat. Rev. Neurosci. *5*, 532–546.

Seitz, R.J., Roland, E., Bohm, C., Greitz, T., and Stone-Elander, S. (1990). Motor learning in man: a positron emission tomographic study. Neuroreport *1*, 57–60.

Shadlen, M.N., and Newsome, W.T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. J. Neurosci. *18*, 3870–3896.

Shadmehr, R., and Mussa-Ivaldi, F.A. (1994). Adaptive representation of dynamics during learning

of a motor task. J. Neurosci. *14*, 3208–3224.

Shadmehr, R., Smith, M.A., and Krakauer, J.W. (2010). Error correction, sensory prediction, and adaptation in motor control. Annu. Rev. Neurosci. *33*, 89–108.

Shen, W., Flajolet, M., Greengard, P., and Surmeier, D.J. (2008). Dichotomous dopaminergic control of striatal synaptic plasticity. Science *321*, 848–851.

Simani, M.C., McGuire, L.M.M., and Sabes, P.N. (2007). Visual-shift adaptation is composed of separable sensory and task-dependent effects. J. Neurophysiol. *98*, 2827–2841.

Smith, M.A., Ghazizadeh, A., and Shadmehr, R. (2006). Interacting adaptive processes with different timescales underlie short-term motor learning. PLoS Biol. *4*, e179.

Stanley, J., and Krakauer, J.W. (2013). Motor skill depends on knowledge of facts. Front Hum Neurosci *7*, 503.

Stein, R.B., Gossen, E.R., and Jones, K.E. (2005). Neuronal variability: noise or part of the signal? Nat. Rev. Neurosci. *6*, 389–397.

Surmeier, D.J., Carrillo-Reid, L., and Bargas, J. (2011). Dopaminergic modulation of striatal neurons, circuits, and assemblies. Neurosci. *198*, 3–18.

Sutton, R.S., and Barto, A.G. (1998). Reinforcement Learning (MIT Press).

Thach, W.T. (1978). Correlation of neural discharge with pattern and force of muscular activity, joint position, and direction of intended next movement in motor cortex and cerebellum. J. Neurophysiol. *41*, 654–676.

Thorndike, E.L. (1898). Animal intelligence: An experimental study of the associative processes in animals. The Psychological Review: Monograph Supplements *2*, i–109.

Thoroughman, K.A., and Shadmehr, R. (2000). Learning of action through adaptive combination of motor primitives. Nature *407*, 742–747.

Todorov, E. (2004). Optimality principles in sensorimotor control. Nat. Neurosci. *7*, 907–915.

Todorov, E. (2005). Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system. Neural Comput *17*, 1084–1108.

Todorov, E., and Jordan, M.I. (2002). Optimal feedback control as a theory of motor coordination. Nat. Neurosci. *5*, 1226–1235.

Tumer, E.C., and Brainard, M.S. (2007). Performance variability enables adaptive plasticity of “crystallized” adult birdsong. Nature *450*, 1240–1244.

Ungerleider, L.G., Doyon, J., and Karni, A. (2002). Imaging brain plasticity during motor skill learning. Neurobiol Learn Mem *78*, 553–564.

Valero-Cuevas, F.J., Venkadesan, M., and Todorov, E. (2009). Structured variability of muscle activations supports the minimal intervention principle of motor control. J. Neurophysiol. *102*, 59–68.

van Beers, R.J., Baraduc, P., and Wolpert, D.M. (2002). Role of uncertainty in sensorimotor control. Philos. Trans. R. Soc. Lond., B, Biol. Sci. *357*, 1137–1145.

van Beers, R.J., Haggard, P., and Wolpert, D.M. (2004). The role of execution noise in movement variability. J. Neurophysiol. *91*, 1050–1063.

Vogels, R., Spileers, W., and Orban, G.A. (1989). The response variability of striate cortical neurons in the behaving monkey. Exp. Brain Res. *77*, 432–436.

Voorn, P., Vanderschuren, L.J.M.J., Groenewegen, H.J., Robbins, T.W., and Pennartz, C.M.A. (2004). Putting a spin on the dorsal-ventral divide of the striatum. Trends Neurosci *27*, 468–474.

Walker, M.P., Brakefield, T., Hobson, J.A., and Stickgold, R. (2003). Dissociable stages of human memory consolidation and reconsolidation. Nature *425*, 616–620.

Wang, S.S., Khiroug, L., and Augustine, G.J. (2000). Quantification of spread of cerebellar long-term depression with chemical two-photon uncaging of glutamate. Proc. Natl. Acad. Sci. USA *97*,

8635–8640.

Welch, J.M., Lu, J., Rodriguiz, R.M., Trotta, N.C., Peça, J., Ding, J.-D., Feliciano, C., Chen, M., Adams, J.P., Luo, J., et al. (2007). Cortico-striatal synaptic defects and OCD-like behaviours in Sapap3-mutant mice. Nature *448*, 894–900.

Welford, A.T. (1971). Fundamentals of Skill (Methuen).

Willingham, D.B. (1998). A neuropsychological theory of motor skill learning. Psychol Rev *105*, 558–584.

Wise, S.P., Moody, S.L., Blomstrom, K.J., and Mitz, A.R. (1998). Changes in motor cortical activity during visuomotor adaptation. Exp. Brain Res. *121*, 285–299.

Wolpert, D.M., Ghahramani, Z., and Jordan, M.I. (1995). Are arm trajectories planned in kinematic or dynamic coordinates? An adaptation study. Exp. Brain Res. *103*, 460–470.

Wu, H.G., Miyamoto, Y.R., Gonzalez Castro, L.N., Olveczky, B.P., and Smith, M.A. (2014). Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. Nat. Neurosci. *17*, 312–321.

Yin, H.H., Knowlton, B.J., and Balleine, B.W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. Eur J Neurosci *19*, 181–189.

Yin, H.H., Mulcare, S.P., Hilário, M.R.F., Clouse, E., Holloway, T., Davis, M.I., Hansson, A.C., Lovinger, D.M., and Costa, R.M. (2009). Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. Nat. Neurosci. *12*, 333–341.

Yin, H.H., Ostlund, S.B., Knowlton, B.J., and Balleine, B.W. (2005). The role of the dorsomedial striatum in instrumental conditioning. Eur J Neurosci *22*, 513–523.

2

CORTICOSTRIATAL DYNAMICS ENCODE THE REFINEMENT OF OUTCOME-RELEVANT VARIABILITY DURING SKILL LEARNING

All data discussed in this chapter is currently under re-revision as the following manuscript: Santos F.J., Jin X., Costa R.M. “Corticostriatal dynamics encode the refinement of outcome-relevant variability during skill learning”.

SUMMARY

Learning to perform a complex motor task requires the optimization of specific behavioral features to cope with task constraints. We show that when mice learn a novel motor paradigm they mainly refine the features that are relevant to obtaining the desired outcome. Animals trained in a progressively more difficult operant task reduced outcome-relevant variability, but not variability in uncorrelated dimensions. Trial-to-trial variability of the activity of motor cortex and striatal projection neurons was higher early in training and subsequently decreased with learning, without concomitant changes in average firing rate. As training progressed, trial-to-trial variability in corticostriatal activity became progressively more correlated with trial-to-trial behavioral variability, but only for the outcome-relevant dimension. Corticostriatal plasticity was required for the reduction in outcome-relevant variability but not for variability in other dimensions. These data suggest that corticostriatal

dynamics encode the refinement of outcome-relevant features during motor learning.

INTRODUCTION

Animals have the ability to learn novel motor skills, which allows them to perform complex patterns of movement and to improve the outcomes of their actions. Acquiring novel skills usually requires exploration of the behavioral space, which is critical for learning (Grunow and Neuringer, 2002; Kao et al., 2005; Miller et al., 2010; Olveczky et al., 2005; Skinner, 1981; Sutton and Barto, 1998; Tumer and Brainard, 2007; Wu et al., 2014). It also requires the selection of the appropriate behavioral features that lead to the desired outcomes (Skinner, 1981). It has been postulated that the motor system can learn complex movements by optimizing motor variability in task-relevant dimensions, correcting only deviations that interfere with the final output of the action (Diedrichsen et al., 2010; Scott, 2004; Todorov and Jordan, 2002; Valero-Cuevas et al., 2009). By optimizing the precision of an action endpoint, for example, humans can perform smooth movements even in the presence of noise (Harris and Wolpert, 1998). Selecting task-relevant features and decreasing task-relevant variability might therefore be a critical component of motor learning (Cohen and Sternad, 2009; Costa, 2011; Franklin and Wolpert, 2008; Shmuelof et al., 2012; Valero-Cuevas et al., 2009).

The reduction of behavior output variability, specifically in relevant domains, suggests that the neural activity underlying the task-relevant behaviors is selected during learning. However, it remains unclear how the dynamics of outcome-relevant variability are encoded at the neural

level. It has been suggested that cortical and basal ganglia circuits are important for the selection of task-relevant features (Barnes et al., 2005; Costa et al., 2004; Jin and Costa, 2010; Kao et al., 2005; Olveczky et al., 2005; Woolley et al., 2014). Consistently, it has been previously shown that the initial stages of learning have increased behavioral (Jin and Costa, 2010; Miller et al., 2010; Tumer and Brainard, 2007) and neuronal (Barnes et al., 2005; Costa et al., 2004) variability, but as specific movements are consolidated, neural variability is reduced in these circuits (Costa et al., 2004; Kao et al., 2005). This suggests that after initial motor and neural exploration, specific behavioral and neuronal patterns are selected and consolidated (Costa, 2011). In this study, we investigated if the dynamics of neural activity in cortical and striatal circuits reflect the changes of behavioral variability in the outcome-relevant domain, and whether corticostriatal plasticity is critical for the refinement of outcome-relevant features.

RESULTS

Outcome-relevant variability is specifically reduced during motor learning

We trained mice to perform a fast lever-pressing task where they were required to press a lever at increasingly higher frequencies, in order to obtain a 20mg food pellet. After introducing the animals (N=20) to the behavioral apparatus and one day of continuous reinforcement, where each lever-press was reinforced, animals were trained intensively with 3 daily sessions for 3 days to perform fast lever presses. In the fast press schedules we introduced a covert target frequency, defined by the inverse of 3 consecutive inter-press intervals (IPI), which increased across sessions from 0Hz to a maximum of 4.5Hz (**Fig. 2.1a**; see

Methods). The rate of lever pressing increased throughout training ($F_{8,152}$=41.34, p<0.0001; **Fig. 2.2a**) and animals rapidly started to organize their behavior in self-paced bouts or sequences of lever presses (**Fig. 2.1e**).

The distribution of the instantaneous lever press frequencies (calculated as the inverse of the inter-press intervals), shows a clear shift from initial sessions, where animals did mostly slow frequency presses (0 - 0.5Hz; but already some higher frequency presses of 0.5 - 4.5Hz and > 4.5Hz), to later sessions where the whole distribution was shifted to faster pressing speeds (**Fig. 2.2c**). Analysis of the log-frequency distributions clearly shows a multimodal distribution with long IPIs (slow press frequencies, <0.5Hz, **Fig. 2.1b and Fig. 2.2d**) representing pauses in pressing or magazine checks. This allowed us to define the bouts or sequences of pressing *a posteriori*, based on the behavioral performance (either by a pause in pressing separated by IPIs higher than 2s,or the occurrence of checking behavior, i.e. magazine pokes between presses; see Methods), independently of the requirements for a specific training session or protocol. This permitted the investigation of the behavior strategies used by each animal in each bout or attempt in relation to the task constraints.

The percentage of lever presses performed within a bout or sequence increased significantly from 56.98 ± 3.98, in the first session of covert target introduction, to 98.26 ± 0.53 in the last training session ($F_{8,152}$=60.22 p<0.0001; **Fig. 2.1c**), and the rate of sequences performed per minute increased with training ($F_{8,152}$=32.23 p<0.0001; **Fig. 2.1d**). The percentage of reinforced sequences tended to decrease as the difficulty of the task increased across sessions, but tended to stabilize or increase when the same target difficulty was repeated in two consecutive sessions ($F_{8,152}$ = 57.31, p<0.0001; **Fig. 2.2b**).

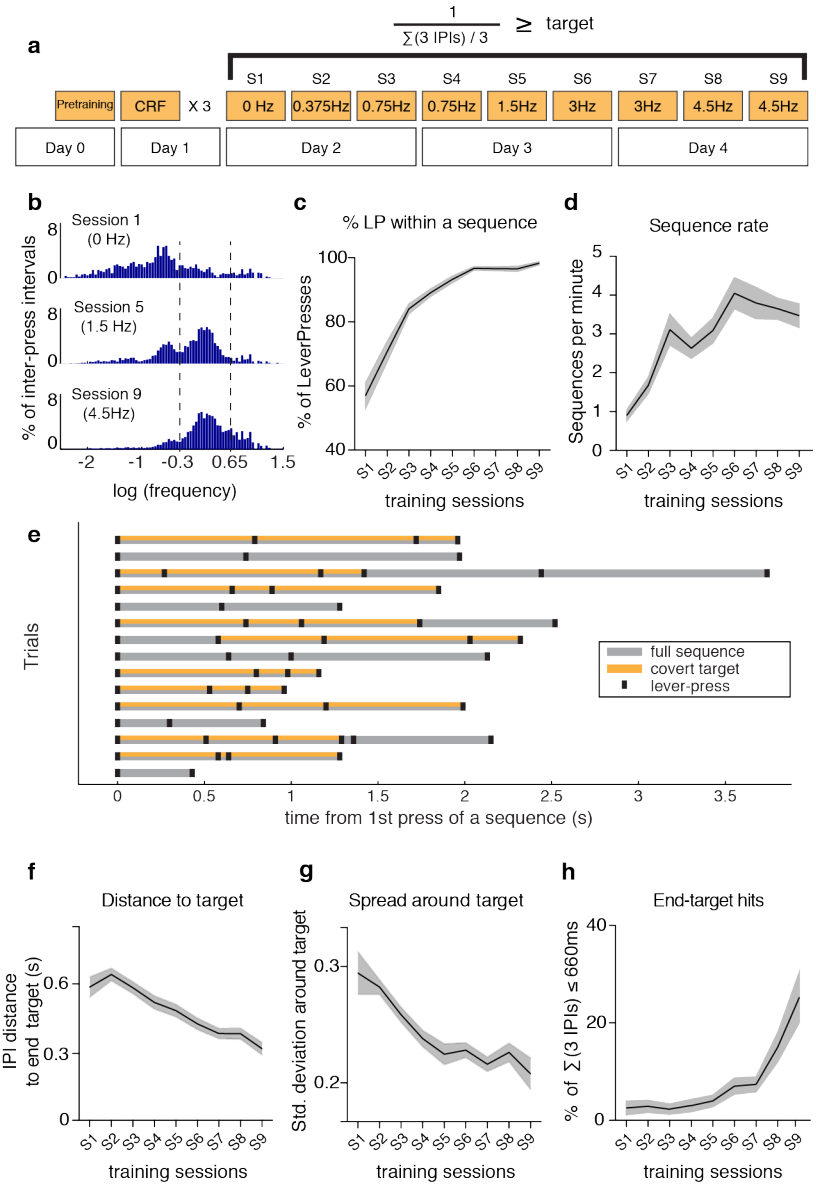


Figure 2.1 | Mice learn a fast lever-pressing task, shaping their behavior to gradually approach the end target. (a) Schematic of the training protocol, starting with magazine habituation and CRF training in the first two days, followed by 3 days of the fast press schedules (S1-S9) where we introduce an increasingly higher covert target, defined as the inverse of the sum of 3 consecutive IPIs. **(b)** Joint distribution of the log(frequency) for all individual IPIs, in the first, middle and last session of the fast press schedules, for all the 20 animals. Vertical dashed lines correspond to the IPI threshold used for sequence definition (IPI=2s, log(freq)=~-0.3) and the final covert target (IPI=3/660ms, log(freq)=~0.65). **(c)** Percentage of lever presses comprised within a

sequence for each training session. **(d)** Number of sequences performed per minute. **(e)** Example of sequences performed by a representative animal, aligned at the time of sequence initiation. Individual lever presses are marked as black ticks, the full sequence duration is shaded in grey and the IPIs that meet the session covert target are shaded in orange **(f)** Distance of the sum of all 3 consecutive IPIs from the final covert target (∑(3IPIs)<660ms, ~4.5Hz) **(g)** Spread of the distance between 3 consecutive IPIs around the final covert target **(h)** Percentage of sequences containing the covert target of the last session (end-target: 3 IPIs < 660ms, ~4.5Hz). Shaded areas correspond to mean ± SEM.

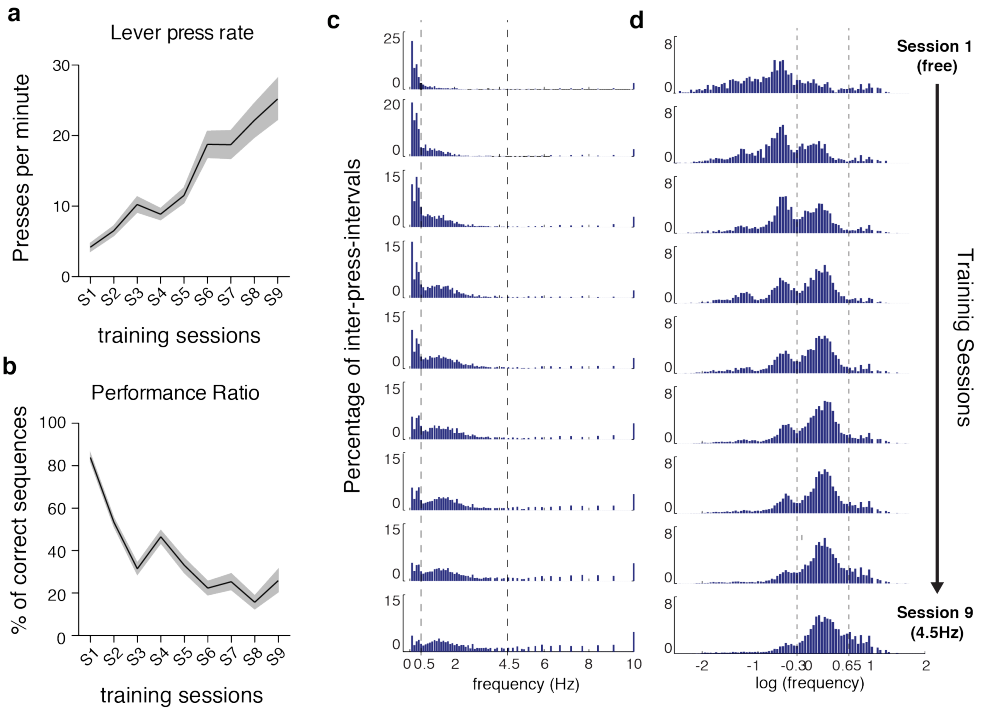


Figure 2.2 | Lever-pressing rate increased and shifted towards higher speeds with training, and performance increased or plateaued when task difficulty did not change in consecutive sessions. (a) Lever presses per minute for the different training sessions ($F_{8,152}$=41.34, p<0.0001) . **(b)** Percentage of reinforced sequences across training sessions ($F_{8,152}$=57.31, p<0.0001, Post hoc comparisons: Fisher's LSD test, (0.75Hz) Session 3 vs Session 4 t_{152}=3.847, p=0.0002; (3Hz) Session 6 vs Session 7 t_{152}=0.7681, p=0.4436; (4.5Hz) Session 8 vs Session 9 t_{152}=2.639, p=0.0092). **(c-d)** Joint distribution of instantaneous lever-press frequencies, defined as the inverse of all the individual IPIs, and log(frequency) across the fast press sessions, of all the 20 animals. Vertical dashed lines correspond to the IPI threshold used for sequence definition (IPI=2s, log(freq)=~-0.3) and the final covert target (IPI=220ms, log(freq)=~0.65). Shaded areas correspond to mean ± SEM.

Importantly, with training, the distance of consecutive IPIs (summed in bins of 3 to match the online criteria for frequency) to the final covert target (target of the last sessions 3 IPIs < 660ms, ~4.5Hz) decreased consistently ($F_{8,152}$=25.76, p<0.0001; **Fig. 2.1f**), indicating that animals shaped their behavior gradually to approach the end target. Not only did the mean distance to the end target decrease, but the spread around the target also decreased ($F_{8,152}$=9.616, p<0.001; calculated as the standard deviation of the distances from the covert target IPI, **Fig. 2.1g**). Consistently, animals gradually increased the percentage of press bouts or sequences that would contain the covert target of the last session (end-target: 3 IPIs < 660ms, ~4.5Hz; $F_{8,152}$=14.15, p<0.0001; **Fig. 2.1h**). These data indicate that animals learned to shape their behavior to get closer to the covert target.

The number of lever-presses in each bout or sequence (sequence length; $F_{8,152}$=37.54, p<0.0001, **Fig. 2.3b**), and the duration of each sequence (sequence duration, $F_{8,152}$=22.69, p<0.0001, **Fig. 2.3c**) increased with training, while the mean frequency of the full sequence decreased slightly (sequence frequency, $F_{8,152}$=2.372, p=0.0195, **Fig. 2.3a**). Importantly, the variability of the behavioral parameters from sequence-to-sequence (measured both by the variance and by the Fano factor, **Fig. 2.3d-i**) was differentially modulated during training. While the variability of sequence frequency decreased significantly throughout training (variance: $F_{8,152}$=4.450, p<0.0001, **Fig. 2.3d**; Fano factor: $F_{8,152}$=5.343, p<0.0001, **Fig. 2.3g)**, the variability of both sequence length and sequence duration did not decrease, and even increased (variance: $F_{8,152}$=13.64 p<0.0001 and $F_{8,152}$=11.15 p<0.0001, **Fig. 2.3e** and **Fig. 2.3f**, respectively**;** Fano factor: $F_{8,152}$=19.65 p<0.0001 and $F_{8,152}$=16.86 p<0.0001; **Fig. 2.3h** and **Fig. 2.3i**, respectively).

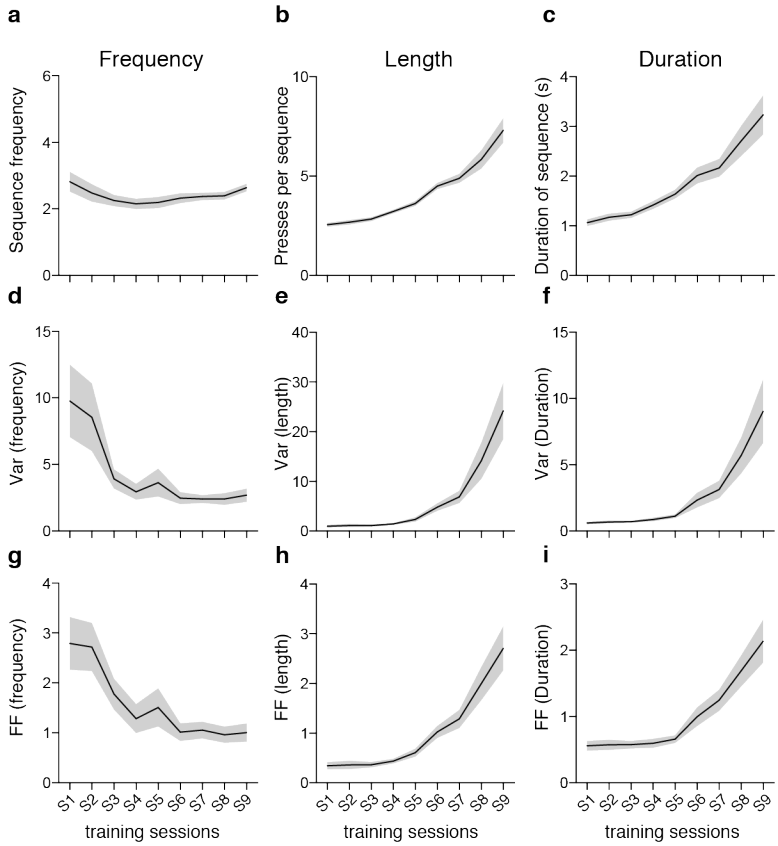
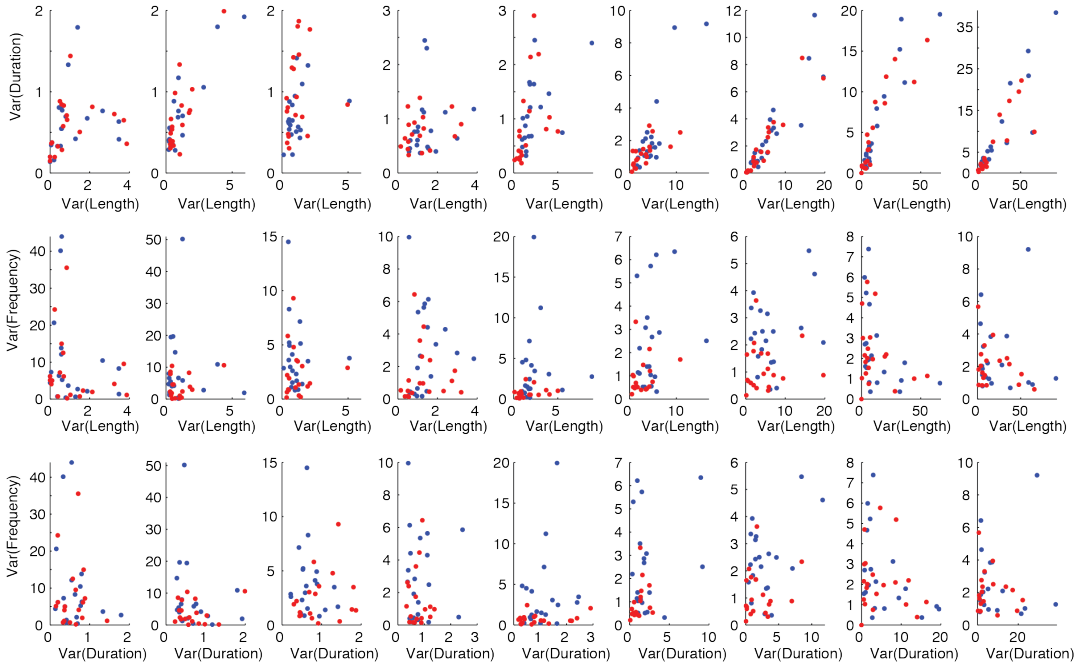


Figure 2.3 | Variability of different behavioral dimensions evolves independently as animals learn a motor task. (a-c) Frequency, number of presses and duration of the full sequences across all the training sessions. **(d-i)** Variability, measured as the variance and Fano factor, for sequence frequency, sequence length and sequence duration. Shaded areas correspond to mean ± SEM.

Furthermore, we observed no correlation between the variance of sequence frequency and the variance of length/duration (**Fig. 2.4**), confirming that variability in frequency and length/duration of a sequence were independently modulated. The decrease of variability in the frequency of pressing from sequence-to-sequence cannot be explained by animals reaching a ceiling in pressing frequency, because the average pressing frequency of whole sequences of presses did not increase with training (it actually decreased slightly). Furthermore,

variability in the frequency domain reached a plateau after session 4 where the target constrains are still rather loose (3 IPIs in less than 4s) and this is a frequency that animals can reach in 78.91 ± 5.09 % of the sequences at the end of training.



r		S1	S2	S3	S4	S5	S6	S7	S8	S9
Length / Duration	All	0.250	0.899	0.342	0.182	0.532	0.869	0.890	0.927	0.908
	Rewarded	0.220	0.913	0.160	0.346	0.558	0.678	0.937	0.902	0.843
Length / Frequency	All	-0.193	-0.042	-0.113	0.041	0.053	0.206	0.320	-0.414	-0.011
	Rewarded	-0.218	0.116	-0.091	0.044	0.188	0.523	0.069	-0.269	-0.293
Duration / Frequency	All	-0.202	-0.197	-0.148	-0.018	0.214	0.197	0.441	-0.425	0.113
	Rewarded	-0.226	0.020	0.114	-0.178	0.250	0.252	0.151	-0.121	-0.143

PVAL		S1	S2	S3	S4	S5	S6	S7	S8	S9
Length / Duration	All	0.287	<0.001	0.141	0.442	0.016	<0.001	<0.001	<0.001	<0.001
	Rewarded	0.351	<0.001	0.500	0.135	0.011	0.001	<0.001	<0.001	<0.001
Length / Frequency	All	0.416	0.861	0.634	0.865	0.826	0.383	0.169	0.069	0.964
	Rewarded	0.356	0.627	0.704	0.855	0.426	0.018	0.772	0.252	0.224
Duration / Frequency	All	0.392	0.405	0.534	0.939	0.366	0.406	0.052	0.062	0.646
	Rewarded	0.339	0.934	0.633	0.453	0.288	0.284	0.525	0.611	0.559

Figure 2.4 | Significant correlation between variance of length and duration, but not between the variance of frequency and length, nor between variance of frequency and duration. Scatter plots of the paired variance for the 3 measured behavior parameters (sequence frequency, length and duration) for each sequence, across nine training sessions. Table shows all the Pearson's linear correlation coefficients and associated p-values for all the comparisons depicted above.

These data suggest that behavioral variability that was relevant to the outcome of the task was specifically refined during training. We therefore analyzed if the variability of each of the behavioral dimensions was different in reinforced vs. non-reinforced sequences (**Fig. 2.5**). We verified that sequences leading to a reinforcer had indeed significantly lower variability in frequency compared to non-reinforced sequences (main effect of reinforcement, $F_{1,38}$=7.608, p=0.0089, **Fig. 2.5d** and $F_{1,38}$=28.34, p<0.0001**, Fig. 2.5g**), but there were no significant differences in the variability of the length/duration between reinforced and non-reinforced sequences (**Fig. 2.5e-f and 2.5h-i**), indicating that outcome-relevant variability was specifically reduced during training.

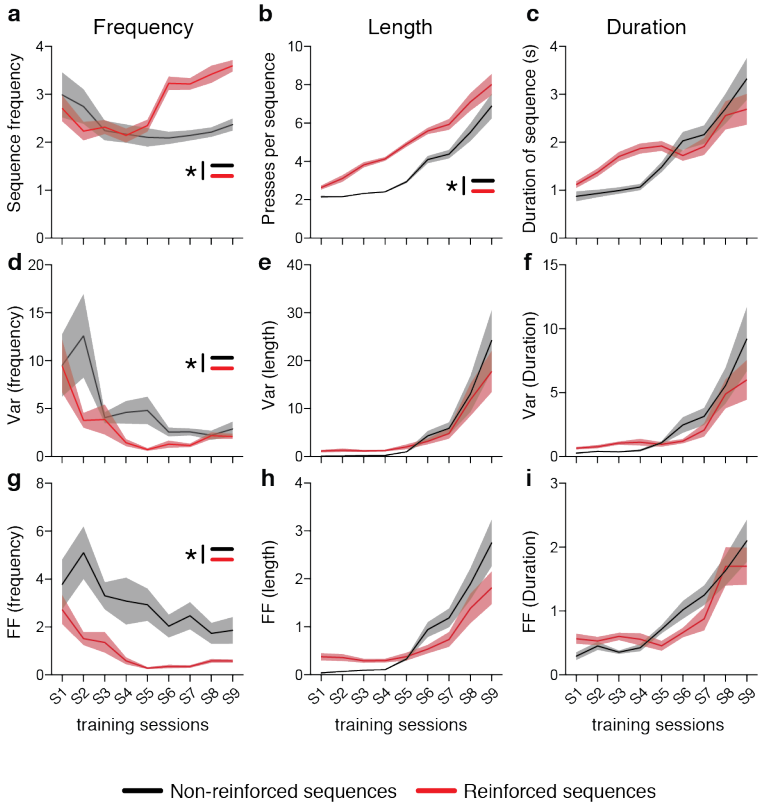
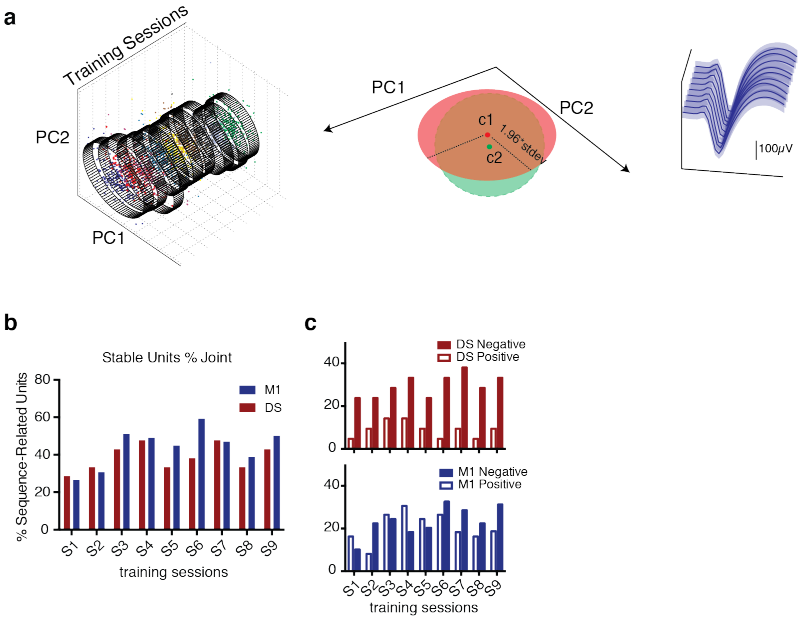


Figure 2.5 | Outcome-relevant variability specifically decreases during training. (a-c) Comparison of frequency, length and duration between reinforced and non-reinforced

sequences. **(d-f)** Variance and **(g-i)** variability, measured as the Fano factor, for the 3 behavior parameters, for reinforced and non-reinforced sequences. Black lines correspond to mean values for non-reinforced sequences. Red lines correspond to mean values for reinforced sequences. Shaded areas correspond to mean ± SEM. * $p<0.05$.

Variability of motor cortex and striatal activity decreases with learning

In order to investigate the dynamics of cortical and striatal circuits during the acquisition and performance of the fast lever pressing task, we continuously recorded extracellular neuronal activity simultaneously in layer 5 of the primary motor cortex (M1), and in the dorsal striatum (DS) of mice during the full duration of training (4 days, N=7 animals, average of 18 M1 units and 10 DS units simultaneously recorded per animal, per session). Non-stop continuous electrophysiological recordings across 4 days encompassing all the sessions of training allowed us to track the activity of a subset of "stable" cells throughout the whole period of training (49 M1 units, 21 DS Units). Putative single-units were isolated based on waveform characteristics, inter-spike intervals (ISI) and clustering statistics using Principal Component Analysis (PCA). Units were considered "stable" if the statistics in PCA space and waveform proprieties did not change significantly across sessions (see Methods and **Fig. 2.6a**).



Stable Units DS		Session								
		S1	S2	S3	S4	S5	S6	S7	S8	S9
Total #		21	21	21	21	21	21	21	21	21
Sequence-Related	Positive	1	2	3	3	2	1	2	1	2
	Negative	5	5	6	7	5	7	8	6	7
	Total	6	7	9	10	7	8	10	7	9

Stable Units M1		Session								
		S1	S2	S3	S4	S5	S6	S7	S8	S9
Total #		49	49	49	49	49	49	49	49	49
Sequence-Related	Positive	8	4	13	15	12	13	9	8	9
	Negative	5	11	12	9	10	16	14	11	15
	Total	13	15	25	24	22	29	23	19	24

Figure 2.6 | Stable cells and sequence-related modulations. (a) Illustration of an example stable cell, with cluster projection using PCA across the training sessions (left), diagram illustrating the criteria for stability of cells across different recording sessions (middle) and the average waveform in each session (right). **(b)** Percentage of sequence-related stable units for both areas across training sessions. **(c)** Stable DS positive (open red bars) and stable negative (filled red bars) sequence-related units, and stable M1 positive (open blue bars) and stable negative (filled blue bars) sequence-related units. Table shows the total number of recorded stable cells and correspondent sequence-related modulations in both M1 and DS for the 7 recorded animals.

We observed, as previously reported, that a high percentage of neurons displayed sequence-related modulations, with some neurons increasing their firing rate and others decreasing firing rate during the performance

of a whole bout or sequence in relation to baseline firing rate levels (Jin and Costa, 2010; Jin et al., 2014). The percentage of neurons showing these modulations increased across both areas as training progressed, suggesting the emergence of sequence-related neuronal ensembles (DS: $F_{8,48}$=3.014, p=0.008, M1: $F_{8,48}$=2.969, p=0.008, **Fig. 2.6a and Fig. 2.7a**). We also confirmed that a high percentage of neurons showed statistically significant modulation related with individual lever-presses, with no change across training (Jin and Costa, 2010) (**Fig. 2.8a**).

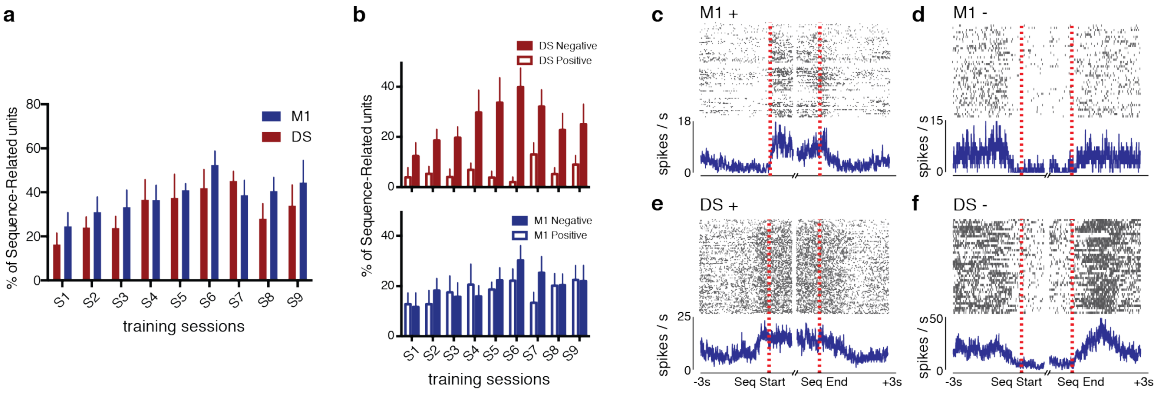


Figure 2.7 | Emergence of sequence-related activity with training. (a) Percentage of sequence-related units for both M1 and DS across training sessions. **(b)** DS positive (open red bars) and negative (filled red bars) sequence-related units, and M1 positive (open blue bars) and negative (filled blue bars) sequence-related units. **(c)** Examples of rasters and PETH aligned to the first and last press of a sequence for positive and negative sequence-related units in both M1 and DS. Error bars correspond to mean ± SEM.

Notably, we found a high sequence-to-sequence variability in the activity of individual neurons (considering each bout or sequence as an attempt; measured by the Fano factor of the firing rate) in the first couple of sessions, that then decreased with training (DS: $F_{8,48}$=2.767 p<0.05; M1: $F_{8,48}$=2.771 p<0.05; **Fig. 2.10a**). These dynamics in neuronal variability were observed during the performance of lever-press sequences, but not during baseline periods (measured from 5 to 2 seconds before the

initiation of each sequence), when the animals were not actively engaged in lever pressing (DS: $F_{8,48}$=1.117 p=0.3324; M1: $F_{8,48}$=1.459 p=0.1973; **Fig. 2.10b**), or during periods flanking the sequence (first press: DS $F_{8,48}$=1.213, p=0.3121; M1 $F_{8,48}$=0.1374, p=0.9971; last press: DS $F_{8,48}$=0.5227, p=0.8335; M1 $F_{8,48}$=0.8677, p=0.5499; **Fig. 2.8b**).

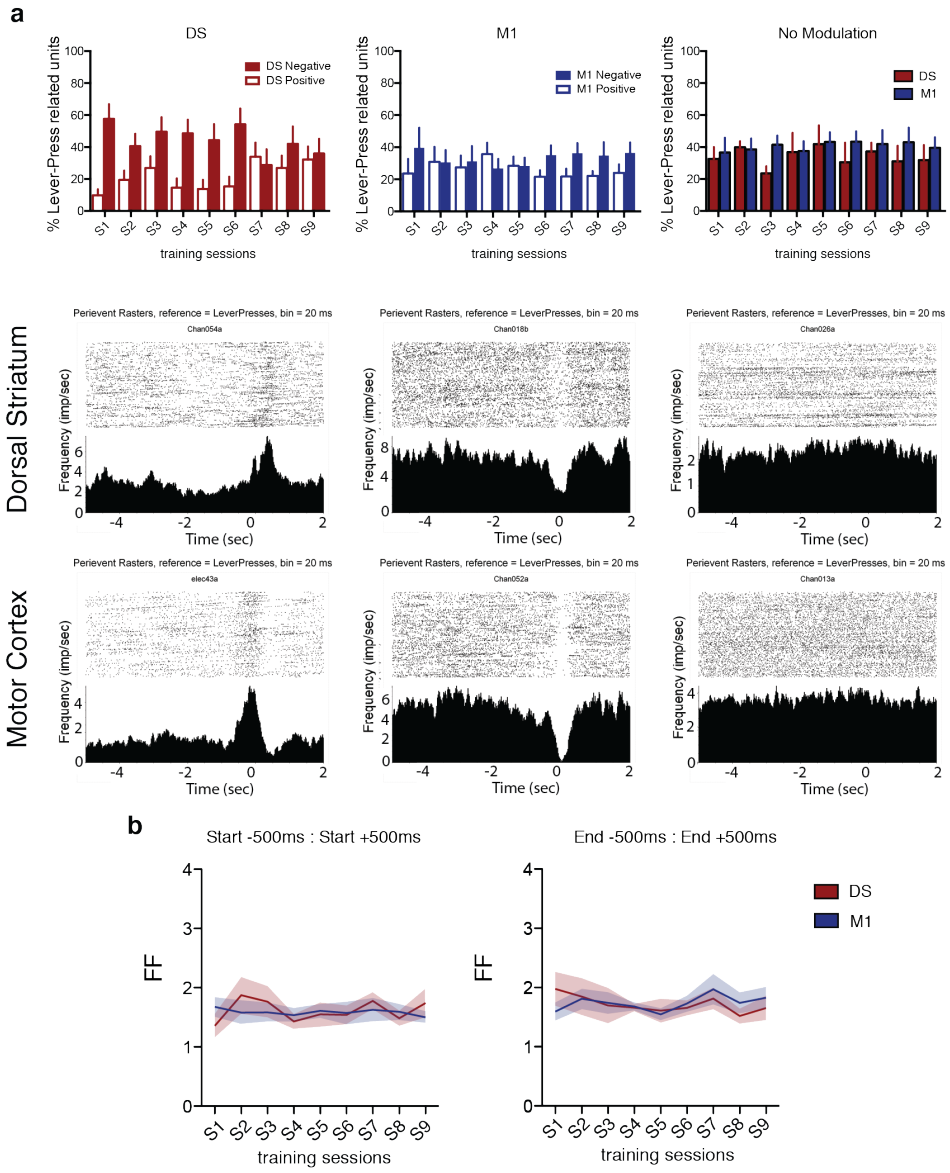


Figure 2.8 | Percentage of individual lever-press related units and variability around the first and last press of a sequence do not change throughout learning. **(a)** Percentage of positive, negative and non-modulated cells with respective examples for both of the recorded areas (DS positive: $F_{8,48}$=1.665, p=0.1317; DS negative: $F_{8,48}$=1.118, p=0.3683; DS no modulation: $F_{8,48}$=0.5048, p=0.8441; M1 positive: $F_{8,48}$=1.363, p=0.2367; M1 negative: $F_{8,48}$=0.5222, p=0.8338; M1 no modulation: $F_{8,48}$=0.2268, p=0.9842) **(b)** Fano factor dynamics calculated for 1s intervals around the first (DS $F_{8,48}$=1.213, p=0.3121; M1 $F_{8,48}$=0.1374, p=0.9971) and last presses (DS $F_{8,48}$=0.5227, p=0.8335; M1 $F_{8,48}$=0.8677, p=0.5499) of a sequence. Shaded areas represent mean ± SEM.

These dynamics do not emerge simply because the behavior is more stereotyped, as variability in behavior decreased in one dimension but increased in others (**Fig. 2**). Furthermore, these dynamics were also observed when restricting the analysis to sequences of matched duration and frequency, where the behavioral dimensions did not change significantly across sessions, indicating that these effects do not emerge from direct changes in the characteristics of each feature of behavior across training (methods and **Fig. 2.9**).

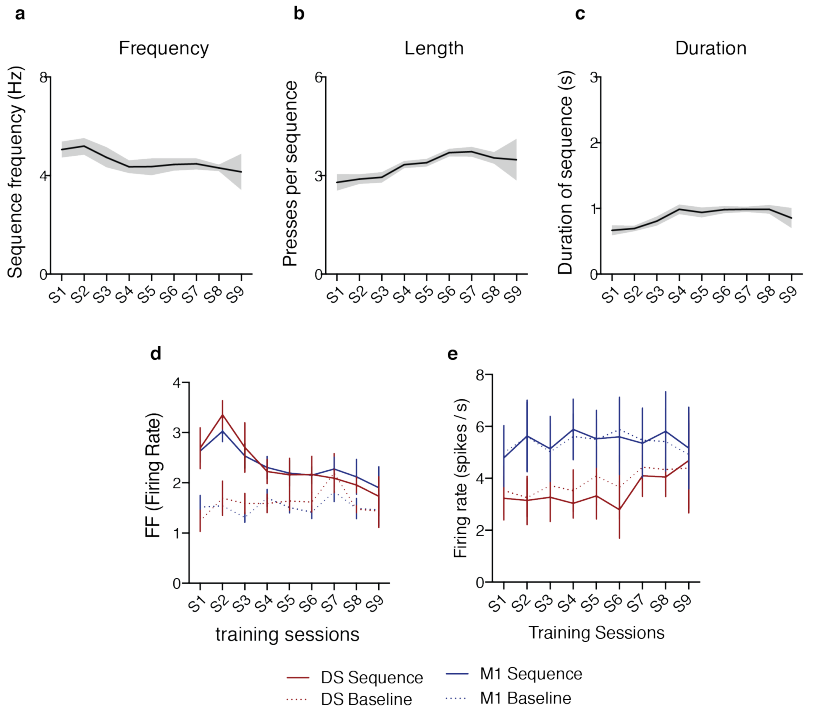


Figure 2.9 | Neuronal variability dynamics are still evident when analysis is restricted to sequences with duration and frequency. (a-c) Frequency, length and duration of matched sequences. **(d)** Neuronal variability, measured as the Fano factor of the firing rate, for sequences of matched duration and frequency, for both recorded areas, during baseline (red dashed line, DS: $F_{8,48}$=1.327, p=0.2532; blue dashed line M1: $F_{8,48}$=0.9318, p=0.4994) and lever-press sequences (red solid line, DS: $F_{8,48}$=2.687, p=0.00158; blue solid line M1: $F_{8,48}$=2.218, p=0.0442). **(e)** Firing rates, for sequences of matched duration and frequency, during baseline and lever-press sequences. Error bars correspond to mean ± SEM.

The decrease in variability was also observed when using exclusively "stable" cells for this analysis (DS: $F_{8,96}$=3.721 p=0.0008; M1 $F_{8,312}$=8.707, p<0.0001; **Fig. 2.10c**), showing that the differences in variability throughout learning were observed in individual cells, and did not represent a shift in the population of neurons recorded across days. Importantly, the average firing rate of individual cells did not change significantly, neither across sessions nor across days ($p > 0.05$ for all conditions, **Fig. 2.10e-h**), suggesting that the reduction in variability was not attributable to overall changes in firing rate, but instead to the selection/refinement of a particular firing patterns related to sequence execution during training.

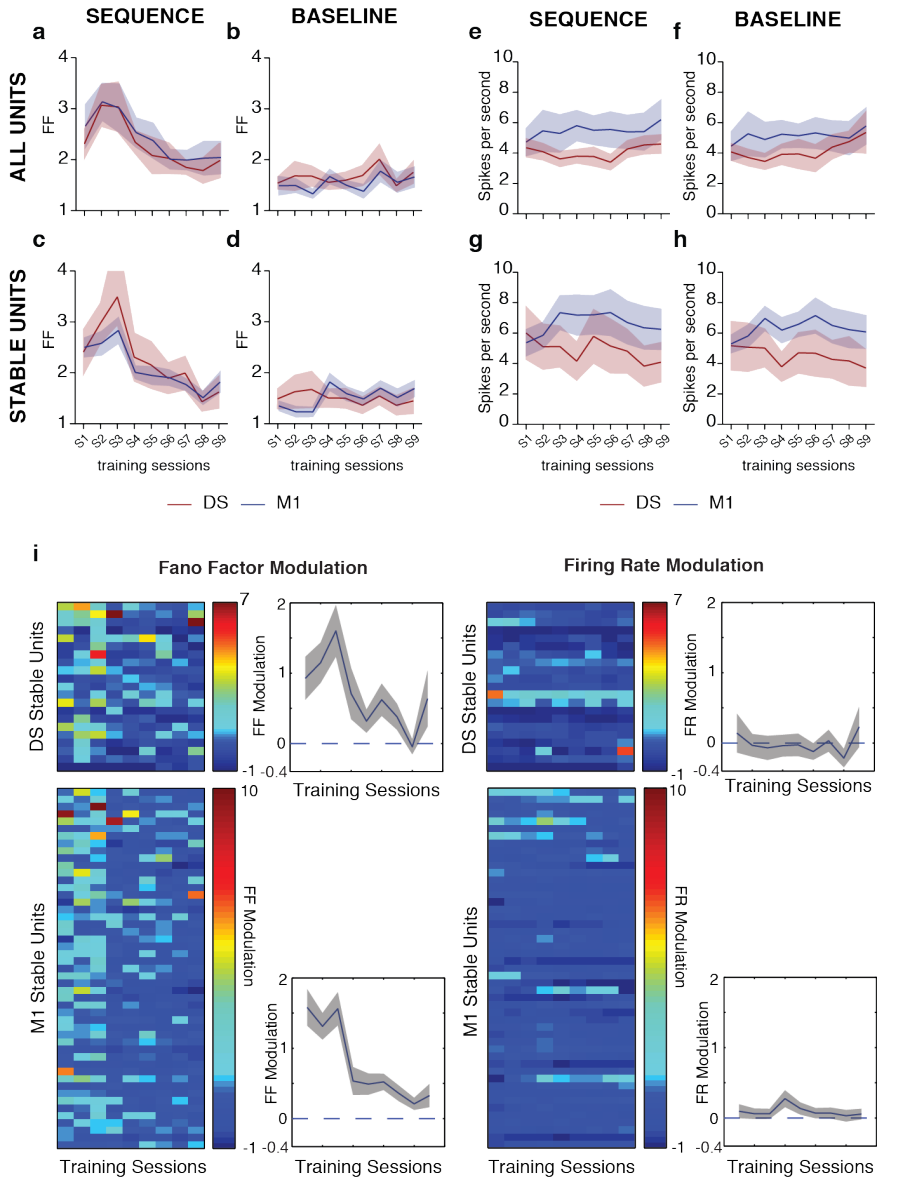


Figure 2.10 | Trial-to-trial variability in corticostriatal circuits decreases throughout training. (a-d) Average neuronal variability (measured as the Fano factor of firing rates) during sequence performance and baseline periods, for all the recorded neuronal units and exclusively for "stable cells", for both M1 (blue traces) and DS (red traces). **(e-h)** Average firing rates during sequence performance and baseline, for all the recorded units and exclusively for stable units, for M1 (blue traces) and DS (red traces). **(i)** Fano factor (FF) and firing rate (FR) modulation relative to baseline values, for individual stable cells recorded across the training sessions within DS (top colorplots) and M1 (bottom colorplots). Right panels depict average modulation. Shaded areas correspond to mean ± SEM.

Further analysis of the dynamics of variability of individual stable cells clearly showed higher variability relative to baseline during the initial sessions (first session DS: W=134, p=0.0107; first session M1: W=1119, p<0.0001), that decreased throughout training until it reached the same levels of baseline at the end of training (last session DS: W=73, p=0.2157; last session M1: W=253, p=0.2121; **Fig. 2.10i**, left panels). Again, average firing rates did not show any significant modulation in relation to baseline throughout the whole period of training (DS: $F_{8,160}$=1.031, p=0.4153; M1: $F_{8,384}$=1.757, p=0.084; **Fig. 2.10i**, right panels).

Corticostriatal variability becomes correlated with behavioral variability exclusively in the outcome-relevant dimension

We next investigated if the reduction in bout-to-bout or sequence-to-sequence variability of neural activity was related to the sequence-to-sequence variability in behavior by analyzing the correlation between neuronal and behavioral dynamics. We re-calculated the Fano factor of the behavioral features and the neuronal activity using a moving average of a reduced number of trials (5) to provide a higher within session resolution of the variability dynamics and therefore permit the correlation of behavioral and neuronal dynamics across training for each animal (**Fig. 2.12a**, see Methods). Analyses of correlations between the variability of the recorded units and the variability of each independent behavior feature revealed a significant increase in correlation between neuronal and behavior variability specifically for the outcome-relevant feature (**Fig. 2.12c**), but not for the remaining features (**Fig. 2.12d-e**). These results show that the decrease in variability in M1 and DS is not just a reflection of a more constrained performance of the movement as

training progresses; variability of the movement decreased in a specific domain but it increased in other domains and the neural variability did not correlate with performance in those domains. Furthermore, no significant correlations were observed between the firing rate of neurons and any of the behavior features (**Fig. 2.11**), indicating again that the observed relationship between neuronal and behavior dynamics was not the reflex of a general increase in correlation between neuronal activity and behavior.

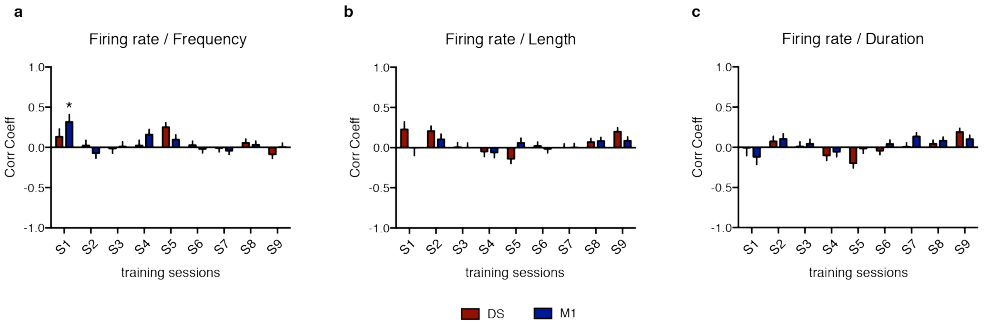


Figure 2.11 | No significant correlation is found between average firing rate and any of the behavior features. Correlation between the average firing rate and **(a)** sequence frequency, **(b)** sequence length, and **(c)** sequence duration. Error bars denote correlation coefficient ± standard error of the correlation. * p>0.05

The data presented above suggested that as training progressed variability in M1 and striatum became more correlated with variability in outcome-relevant behavior. This suggests that neural variability in M1 and striatum could also become more coupled with training. We verified that at the onset of training the sequence-to-sequence variability of neural activity in DS and M1 in each animal was not correlated. However, a strong correlation between the variability in DS and M1 rapidly emerged during training ($p<0.05$ for all except the first training session, **Fig. 2.12b**), suggesting that as outcome-related behavioral variability is refined, neural variability in M1 and striatum becomes coupled.

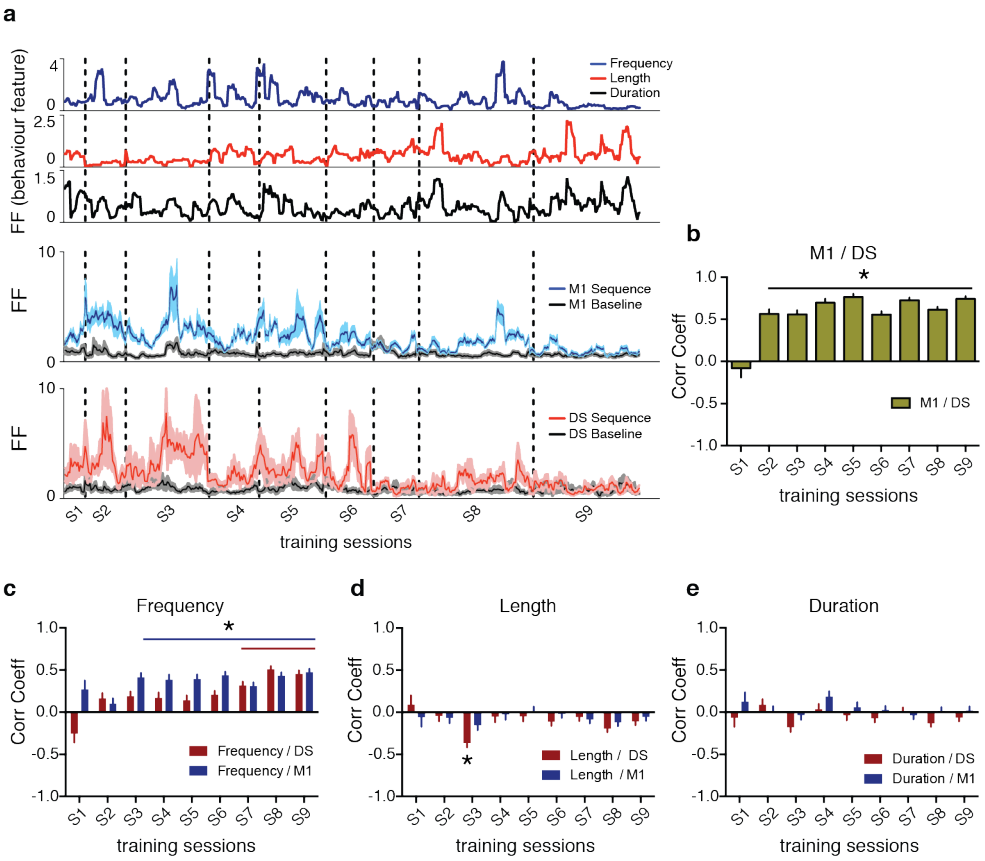


Figure 2.12 | Correlations between corticostriatal and behavioral variability emerge specifically for outcome-relevant features. (a) Example traces from a single animal representing variability, calculated as the Fano factor, using a moving window of 5 consecutive trials shifted by one for sequence frequency (top blue trace), sequence length (top red trace), sequence duration (top black trace), M1 units firing rate during sequences (bottom blue trace) and baseline (dashed grey trace), and DS units firing rate during sequences (bottom red trace) and baseline (dashed grey trace). Vertical dashed lines represent separation of different training sessions. Shaded areas correspond to mean ± SEM. **(b)** Correlation between the variability (FF) in M1 and DS. **(c-e)** Correlation between variability traces from neuronal firing rates in M1 (blue bars) or DS (red bars), and variability of sequence frequency, length and duration. Error bars denote r ± SE_r. * $p<0.05$.

Corticostriatal plasticity is required for the reduction in outcome-relevant variability

The results presented above show that a coupled reduction in corticostriatal variability accompanies the reduction in variability of the outcome-relevant behavioral feature, but not the other features, suggesting that corticostriatal plasticity is necessary to select the appropriate motor features and hence reduce variability in the outcome-relevant domain. Therefore we decided to directly investigate if the observed reduction in outcome-relevant variability is dependent on corticostriatal plasticity by using mutant mice with disrupted NMDA receptors specifically at glutamatergic synapses of striatal projection neurons (*RGS9-Cre / NMDAR1-loxP*), which have impaired corticostriatal plasticity (Dang et al., 2006), and control littermates. Mutant animals had more difficulty learning the task, so we adapted the training protocol to one session per day for both mutant and littermate controls (and repeated sessions when needed), in order to achieve comparable performance levels (see Methods and **Fig. 2.13a**).

As expected, the average distance to target (Controls: $p=0.0450$, $t_5=2.657$, **Fig. 2.13b**) and spread around the target (Controls: $p=0.0179$, $t_5=3.466$, **Fig. 2.13c**) decreased in controls. However, neither of these measures changed with training in *RGS9-Cre / NMDAR1-loxP* mutants (Mutants: $p=0.3535$, $t_6=1.005$; and $p=0.2817$, $t_6=1.183$, respectively).

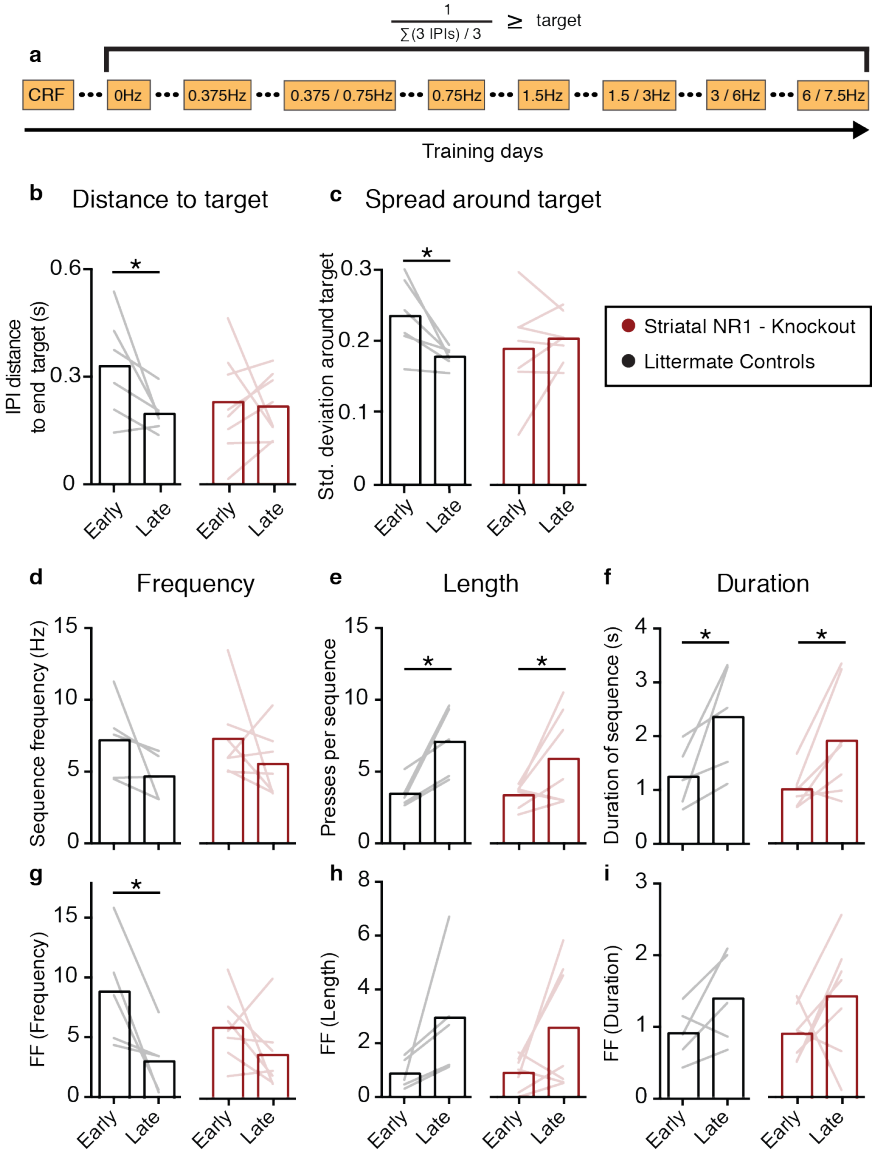


Figure 2.13 | Corticostriatal plasticity is necessary specifically for the reduction in outcome-relevant variability. (a) Schematic of the adapted training sessions for mutant animals and littermate controls. Animals would remain in the same training session until reaching a stable performance. **(b)** Distance of the sum of all 3 consecutive IPIs from the final covert target (∑(3IPIs)<660ms, ~4.5Hz) in *RGS9-NR1* mutants and littermate controls **(c)** Spread of the distance between 3 consecutive IPIs around the final covert target. **(d-i)** Behavior parameters and variability, measured as the Fano factor, during early and late training sessions in *RGS9-NR1* mutants and littermate controls groups. Bars correspond to mean, with data from individual animals plotted on the background (red lines: *RGS9-NR1* KO; black lines: littermate controls).

In general, no significant difference was observed for any of the behavior features between the two groups of animals. However, planned comparisons did show that *RGS9-Cre / NMDAR1-loxP* mutants did not decrease outcome-relevant variability during training, in contrast to littermate controls which did (significant main effect of training time: $F_{1,10}$=10.13, p=0.009; Posthocs: Mutant group: t_{10}=1.38, p=0.1964; Control group: t_{10}=3.00, p=0.0134). Importantly, no differences in the modulation of outcome-irrelevant features were observed between the two groups (no significant main effect for genotype in both conditions; Length FF: $F_{1,10}$=0.06, p=0.818; Duration FF: $F_{1,10}$=0.02, p=0.887) (**Fig. 2.13d-i**). These data suggest that corticostriatal plasticity is required for the reduction in variability of the outcome-relevant behavioral features.

DISCUSSION

The present results show that the acquisition of a novel motor paradigm involving the performance of implicit fast sequences of the same action results in differential modulation of the variability of different behavioral features. Studies in humans have shown a reduction in variability during skill acquisition, with differential modulation of the different components of a task space (Cohen and Sternad, 2009; Müller and Sternad, 2004). According to the “minimum intervention principle” (Todorov and Jordan, 2002), motor skill learning follows some proprieties of optimal feedback control, where instead of a reducing variability in all the components of the task, the motor system corrects only for deviations that interfere with the goals of the action, optimizing task-relevant features while allowing for variability in features that are not relevant for outcome of the task. In this study, we used a paradigm where mice have to perform specific patterns of the same action to obtain an outcome, to show that this

principle applies to sequences or bouts of actions. Contrary to tasks that use sequences of different actions (heterogeneous sequences), where order matters and explicit strategies can be more easily implemented (Ghilardi et al., 2009; Tanji, 2001), in this task only the specific temporal arrangement or the number of presses (and correlated duration) can be shaped in each sequence. Using this paradigm, we show specific refinement of outcome-relevant variability in the performance of a sequence of actions (in this case frequency), but not of uncorrelated variability in other dimensions of the sequence (length/duration).

As per the task structure, animals can use different behavioral strategies to solve the task. It appears that mice increased the number of presses per sequence and the sequence-to-sequence variability in number of presses while concomitantly increasing the probability of getting to the covert target and decreasing the sequence-to-sequence variability in press frequency. Because sequences or bouts were not defined by the task requirements but by the statistics of the behavior (pausing or checking the magazine), many sequences that contain hidden targets do not show a change in average frequency of the whole sequence. Still, the variability in the frequency at which these sequences are executed became less variable with training, and this is especially true for sequences that lead to a successful outcome. Furthermore, the number of sequences with a correct end target increased with training and the distance to end target decreased with training, indicating that mice implicitly learned to shape their behavior to get closer to the target. As mice shape their behavior to be closer to the implicit target, they specifically reduce outcome-relevant variability, suggesting that the behavioral patterns that are more relevant for obtaining outcomes are selected during training.

In parallel with the behavior re-organization, we observed initial high sequence-to-sequence variability of neuronal activity in corticostriatal circuits that decreased with training. Variability in the spike patterns of individual neurons and populations of neurons may be the bases for a process of behavioral exploration (or trial) (Kao et al., 2005; Mandelblat-Cerf et al., 2009; Olveczky et al., 2005), while a decrease in neural variability may reflect a process of selection of specific patterns of neural activity that lead to specific behavioral outputs (Costa et al., 2004; Fee and Goldberg, 2011; Kao et al., 2005). It has been suggested that a decrease in corticostriatal variability as a motor task is learned (Barnes et al., 2005; Costa et al., 2004) could correspond to the process of selection and consolidation of specific motor patterns (Costa, 2011). Here, we show that this decrease in neural variability in corticostriatal circuits correlates specifically with the decrease in variability in the outcome-relevant domain. These data provocatively suggest that activity in motor cortex - striatum circuits do not necessarily relate to the optimization of all movement features, but rather to the outcome-relevant features of the movement. Furthermore, we show that variability in cortical and striatal circuits decreases with learning, and that corticostriatal plasticity is important for the refinement of outcome-relevant features. Our data therefore suggests an important role for corticostriatal dynamics in selecting the appropriate implicit behavioral patterns that lead to desired outcomes. Still, corticostriatal variability did not correlate with changes in non-outcome-relevant variability, further suggesting that corticostriatal circuits are necessary for the selection of the variability that is relevant for outcome obtainment (Costa, 2011), but not necessarily for its generation (Goldberg and Fee, 2011). Although in this study we don't investigate the mechanisms underlying the generation of variability, several studies have suggested that the basal ganglia, dopaminergic system, specific cortical circuits, or cerebellar

circuits could subserve this function (Costa, 2011; Costa et al., 2006; Fee and Goldberg, 2011; Leblois et al., 2010; Olveczky et al., 2005; Shmuelof and Krakauer, 2011; Woolley et al., 2014).

Taken together, these findings show that the selection and refinement of specific features that lead to desired outcomes during motor learning is accompanied by the selection of neural patterns in corticostriatal circuits that correlate specifically with the behaviorally relevant dimensions; plasticity in these circuits seems to be critical for the selection of the outcome-relevant features.

METHODS

Animals

All experimental procedures were carried in accordance to the ethics committee guidelines of the Champalimaud Foundation and Instituto Gulbenkian de Ciência, and with approval of the Portuguese DGAV. Experimental procedures were carried out using twenty male, 3 to 5 month old C57BL6/J mice. From these, thirteen animals were used exclusively for behavioral training while the remaining seven underwent microelectrode array implantation for neuronal data recordings. Animals were maintained on a light-dark cycle of 12h:12h starting at 7AM. All experiments were done during the light cycle. Mice were housed in groups of 4 animals prior to surgery and individually after the electrodes were implanted. Three to 6 months old *RGS9L-Cre/Nr1f/f* homozygous mice (N=7) and *Cre* negative littermate controls (N=5) were used for the mutant mouse behavioral experiments.

Surgery and *in vivo* extracellular recordings

Seven C57Bl6/J mice were implanted bilaterally with two micro-electrode arrays (2x8), 35-50 µm tungsten electrodes with micro-polished tips. One array targeted the primary motor cortex (M1, layer 5) while the second was targeting the dorsal striatum (DS, sensorimotor area that receives projections from the same area in M1). Craniotomies and electrode array positioning were done according to coordinates from the Mouse Brain Atlas (Paxinos and Franklin, 2004). M1 array was placed 1mm rostral and 1.6mm lateral from bregma, and lowered ~ 1mm from the surface of the brain, in an area corresponding to the forelimb representation (Tennant et al., 2011). DS array was placed 0.5mm rostral and 2.1mm lateral from bregma, and lowered ~ 2.3mm from the surface of the brain. Electrodes were manually lowered at slow rates while constantly monitoring neural activity in all the channels in order to control for proper electrode function and correct positioning. Final verification of electrode position was done after all the experimental procedures were finished, by perfusing animals with PFA and histological confirmation of Nissl stained 70µm brain slices (**Fig. 2.14**).

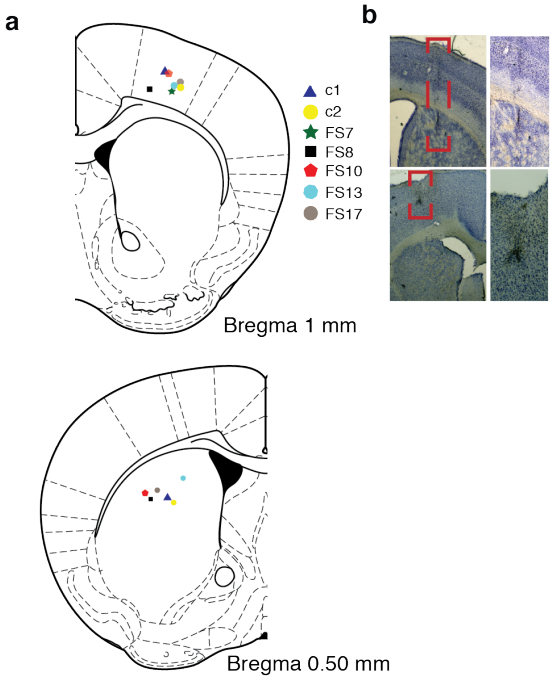


Figure 2.14 | Histological confirmation of electrode tip position used for continuous recording of neuronal activity in corticostriatal circuits. (a) Depiction of electrode array tip localization for motor cortex (top) and dorsal striatum (bottom) for each individual animal. **(b)** Example coronal brain slice magnification using cresyl violet staining for confirmation of electrodes position. Atlas adapted from Paxinos & Franklin (2004).

After surgery animals were allowed to recover for at least two weeks before starting any other experimental procedure. Single and multi unit activity was recorded using Blackrock Microsystems Neural Signal Processor, allowing for online sorting of identified units. Further offline sorting of selected units was done using Plexon Offline Sorter v3 (Plexon, Inc), based on waveform characteristics, inter-spike intervals and PCA clustering. Units stability was assessed from waveforms and PCA cluster proprieties. For PCA cluster comparison data from all the training sessions was pooled together to calculate common eigen vectors. Data from individual sessions was then projected into this common PC space, allowing us to determine cluster centroids and dispersion for each session. Clusters were considered stable whenever the centroid in a given session was comprised within the interval of the centroid of the previous session ± 1.96 * standard deviation of the cluster, in the first two principal components (**Fig. 2.6a** for a graphical representation of this criteria)

Behavioral training

Animals were trained using operant chambers (MedAssociates, Inc) placed inside sound attenuating boxes. A retractable lever was extruded in the beginning of each session, simultaneous to the onset of a light. Animals were required to perform a sequence of presses at a minimum frequency in order to obtain a 20mg food pellet (Bio-serv). 24h before the first training session animals were placed under a food restriction schedule. Body weight was constantly monitored in order to be kept above 85% of the initial weight. In order to facilitate learning, animals were initially exposed to one session of magazine training were food pellets would be available on a random time schedule, and to three sessions of continuous reinforcement schedule (CRF) one day before training, where single lever presses would be reinforced. After that, the

requirements of the task changed and animals were reinforced if they performed an implicit or covert target of consecutive presses, where the frequency, defined by the inverse of 3 consecutive inter-press intervals (IPI), increased with training. On the first session there was no minimum frequency covert target, meaning that any consecutive 3 IPIs would lead to reinforcement. In consecutive sessions the minimum frequency that would lead to reinforcement was increased or maintained in the following order: 0.375Hz, 0.75Hz, 0.75Hz, 1.5Hz, 3Hz, 3Hz, 4.5Hz and 4.5Hz. This constant increase in the minimum frequency of the covert target forced the animals to systematically adapt to the task requirements and perform faster sequences of presses from session to session. The training protocol for mutant animals and littermate controls was adapted due to difficulties learning the task, to one daily session and using automatic progressive schedules once a minimum number of reinforcements (30 or 10) was achieved. (**Table 2.1** for performance summary).

Table 2.1 | Training protocol and respective number of animals reaching performance criteria for the RGS9 – NR1 mutants and littermate controls.

Training Protocol		free	0.375Hz	0.375 / 0.75Hz (30 reinf)	0.75Hz	1.5Hz	1.5 / 3Hz (30 reinf)	3 / 6Hz (10reinf)	6 / 7.5Hz (10 reinf)
# of subjects reaching criteria	NR1 - KO	7	7	6	6	5	4	1	1
	Controls	5	5	5	5	5	5	5	2

Sequences of lever presses

Sequences of presses were differentiated based on inter-press interval (IPI) and occurrence of a magazine head entry. An IPI > 2 seconds (determined based on the distribution of IPIs) or a head-entry were used to define the bouts or sequences of presses. The 2 seconds cutoff was

determined from the joint distribution of the instantaneous IPIs (and the corresponding log distribution) from all the animals, by determining the valley between the two main peaks of IPIs (**Fig. 2.2j-k**). Length of each sequence was defined as the number of press events in each sequence. Duration of each sequence was defined as the time between the first and the last press event. Frequency of each sequence was defined as the inverse of the average inter-press interval of each sequence. For the matched sequences analysis, sequences were selected simultaneously for frequency and duration, based on a short duration (0.2-2s) and high frequency (>2Hz)

Task-related neurons

Neural activity was averaged in 20-ms bins, shifted by 1 ms, and averaged across trials to construct the peri-event histogram (PETH). Data from the PETH from 5000 to 2000 ms before lever press were considered as baseline activity. A positive modulation in firing rate was defined if at least 20 consecutive bins had firing rate larger than a threshold of 99 % above baseline activity, and a negative modulation of firing rate was defined if at least 20 consecutive bins had a firing rate smaller than a threshold of 95 % below baseline activity (Belova et al., 2007). Paired t-tests between baseline firing rate and sequence firing rate were used to classify individual neurons as sequence-related.

Analysis and statistics

All analysis were done in Matlab (The Mathworks, Inc) using custom written programs, or using GraphPad Prism (GraphPad Software, Inc). Normality was verified for all tests using the D'Agostino-Pearson omnibus normality test, or the Kolmogorov-Smirnov test when sample size was too small. Repeated measures ANOVA were used to evaluate

changes in behavior and neuronal features. Paired t-tests were used to evaluate differences in percentage of lever-presses. Increases in FF modulation were assessed by the Wilcoxon Rank Signed test. Repeated measures two-way ANOVA was used to verify the general effect of the *RGS9-NR1* mutants experiment. Sample sizes were calculated based on α= 0.05 and power of 0.7. Trial-to-trial variability of neuronal and behavior data was assessed using Fano factor. We calculate the Fano factor of individual units by dividing the variance of firing rates across all the trials of a session (**Fig. 2.10a-d**,**i**) by the mean over those trials. Fano factor and firing rate modulations for individual stable cells (**Fig. 2.10i**) were calculated as the ratio between the difference of values for sequence and baseline and the values during baseline (Fano factor: $[FF_{sequence}-FF_{baseline}] / FF_{baseline}$; firing rate: $[FR_{sequence}-FR_{baseline}] / FR_{baseline}$). Fano factor of the behavioral features was calculated by dividing the variance in the individual features by the mean of the feature for all the trials (**Fig. 2.3g-i and 2.5g-i**). To establish correlations between the variability of the neuronal data and the variability of the behavior (**Fig. 2.12**), Fano factors were calculated using five consecutive trials, allowing us to increase the resolution of the variability measures. Correlations between neuronal and behavior data were evaluated using Pearson's linear correlations. To avoid correlations bias due to sample size, statistical significance of all the correlations was assessed using the significance criteria for the session with smaller size. Within animal correlations averaged using Fisher's z transformation (Silver and Dunlap, 1987) returned similar results to grouped correlations for all the tested conditions (data not shown).

ACKNOWLEDGMENTS

We thank V. Paixão and A. Gomez-Marin for valuable comments on the manuscript and A. Vaz for colony management. This research was supported by the INDP Graduate Programme and a FCT fellowship to F.J.S., and an European Research Council Starting Grant and a HHM1 International Early Career Scientist Grant to R.M.C.

AUTHOR CONTRIBUTIONS

F.J.S. performed the behavioral and recording experiments and analyses. F.J.S. and R.M.C. designed the experiments and wrote the manuscript. X.J. performed the RGS9-Cre / NMDAR1-loxP experiment and F.J.S. analysed the data.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

REFERENCES

Barnes, T.D., Kubota, Y., Hu, D., Jin, D.Z., and Graybiel, A.M. (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. Nature *437*, 1158–1161.

Belova, M.A., Paton, J.J., Morrison, S.E., and Salzman, C.D. (2007). Expectation Modulates Neural Responses to Pleasant and Aversive Stimuli in Primate Amygdala. Neuron *55*, 970–984.

Cohen, R.G., and Sternad, D. (2009). Variability in motor learning: relocating, channeling and reducing noise. Exp. Brain Res. *193*, 69–83.

Costa, R.M. (2011). A selectionist account of de novo action learning. Curr. Opin. Neurobiol. *21*, 579–586.

Costa, R.M., Cohen, D., and Nicolelis, M.A.L. (2004). Differential corticostriatal plasticity during fast and slow motor skill learning in mice. Curr. Biol. *14*, 1124–1134.

Costa, R.M., Lin, S.-C., Sotnikova, T.D., Cyr, M., Gainetdinov, R.R., Caron, M.G., and Nicolelis, M.A.L. (2006). Rapid alterations in corticostriatal ensemble coordination during acute dopamine-dependent motor dysfunction. Neuron *52*, 359–369.

Dang, M.T., Yokoi, F., Yin, H.H., Lovinger, D.M., Wang, Y., and Li, Y. (2006). Disrupted motor learning and long-term synaptic plasticity in mice lacking NMDAR1 in the striatum. Proc. Natl. Acad. Sci. USA *103*, 15254–15259.

Diedrichsen, J., Shadmehr, R., and Ivry, R.B. (2010). The coordination of movement: optimal

feedback control and beyond. Trends Cogn. Sci. *14*, 31–39.

Fee, M.S., and Goldberg, J.H. (2011). A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. Neurosci. *198*, 152–170.

Franklin, D.W., and Wolpert, D.M. (2008). Specificity of reflex adaptation for task-relevant variability. J. Neurosci. *28*, 14165–14175.

Ghilardi, M.F., Moisello, C., Silvestri, G., Ghez, C., and Krakauer, J.W. (2009). Learning of a sequential motor skill comprises explicit and implicit components that consolidate differently. J. Neurophysiol. *101*, 2218–2229.

Goldberg, J.H., and Fee, M.S. (2011). Vocal babbling in songbirds requires the basal ganglia-recipient motor thalamus but not the basal ganglia. J. Neurophysiol. *105*, 2729–2739.

Grunow, A., and Neuringer, A. (2002). Learning to vary and varying to learn. Psychon Bull Rev *9*, 250–258.

Harris, C.M., and Wolpert, D.M. (1998). Signal-dependent noise determines motor planning. Nature *394*, 780–784.

Jin, X., and Costa, R.M. (2010). Start/stop signals emerge in nigrostriatal circuits during sequence learning. Nature *466*, 457–462.

Jin, X., Tecuapetla, F., and Costa, R.M. (2014). Basal ganglia subcircuits distinctively encode the parsing and concatenation of action sequences. Nat. Neurosci. *17*, 423–430.

Kao, M.H., Doupe, A.J., and Brainard, M.S. (2005). Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. Nature *433*, 638–643.

Leblois, A., Wendel, B.J., and Perkel, D.J. (2010). Striatal dopamine modulates basal ganglia output and regulates social context-dependent behavioral variability through D1 receptors. J. Neurosci. *30*, 5730–5743.

Mandelblat-Cerf, Y., Paz, R., and Vaadia, E. (2009). Trial-to-trial variability of single cells in motor cortices is dynamically modified during visuomotor adaptation. J. Neurosci. *29*, 15053–15062.

Miller, J.E., Hilliard, A.T., and White, S.A. (2010). Song practice promotes acute vocal variability at a key stage of sensorimotor learning. PLoS ONE *5*, e8592.

Müller, H., and Sternad, D. (2004). Decomposition of variability in the execution of goal-oriented tasks: three components of skill improvement. J Exp Psychol Hum Percept Perform *30*, 212–233.

Olveczky, B.P., Andalman, A.S., and Fee, M.S. (2005). Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. PLoS Biol. *3*, e153.

Paxinos, G., and Franklin, K.B.J. (2004). The Mouse Brain in Stereotaxic Coordinates (Gulf Professional Publishing).

Scott, S.H. (2004). Optimal feedback control and the neural basis of volitional motor control. Nat. Rev. Neurosci. *5*, 532–546.

Shmuelof, L., and Krakauer, J.W. (2011). Are we ready for a natural history of motor learning? Neuron *72*, 469–476.

Shmuelof, L., Krakauer, J.W., and Mazzoni, P. (2012). How is a motor skill learned? Change and invariance at the levels of task success and trajectory control. J. Neurophysiol. *108*, 578–594.

Silver, N.C., and Dunlap, W.P. (1987). Averaging correlation coefficients: Should Fisher's z transformation be used? J. Appl. Psychol. *72*, 146–148.

Skinner, B.F. (1981). Selection by consequences. Science *213*, 501–504.

Sutton, R.S., and Barto, A.G. (1998). Reinforcement Learning (MIT Press).

Tanji, J. (2001). Sequential organization of multiple movements: involvement of cortical motor areas. Annu. Rev. Neurosci. *24*, 631–651.

Tennant, K.A., Adkins, D.L., Donlan, N.A., Asay, A.L., Thomas, N., Kleim, J.A., and Jones, T.A.

(2011). The organization of the forelimb representation of the C57BL/6 mouse motor cortex as defined by intracortical microstimulation and cytoarchitecture. Cereb Cortex *21*, 865–876.

Todorov, E., and Jordan, M.I. (2002). Optimal feedback control as a theory of motor coordination. Nat. Neurosci. *5*, 1226–1235.

Tumer, E.C., and Brainard, M.S. (2007). Performance variability enables adaptive plasticity of "crystallized" adult birdsong. Nature *450*, 1240–1244.

Valero-Cuevas, F.J., Venkadesan, M., and Todorov, E. (2009). Structured variability of muscle activations supports the minimal intervention principle of motor control. J. Neurophysiol. *102*, 59–68.

Woolley, S.C., Rajan, R., Joshua, M., and Doupe, A.J. (2014). Emergence of context-dependent variability across a basal ganglia network. Neuron *82*, 208–223.

Wu, H.G., Miyamoto, Y.R., Gonzalez Castro, L.N., Olveczky, B.P., and Smith, M.A. (2014). Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. Nat. Neurosci. *17*, 312–321.

3

CLOSED-LOOP REINFORCEMENT OF MOTOR CORTEX ACTIVITY LEADS TO THE SELECTION OF SPECIFIC SPATIOTEMPORAL ACTIVITY PATTERNS

All data discussed in this chapter is currently in preparation as the following manuscript: Santos F.J., Costa R.M. “Closed-loop reinforcement of motor cortex activity leads to the selection of specific spatiotemporal activity patterns”.

SUMMARY

Action selection allows animals to select from exploratory motor programs in order to optimize the cost function of their behavior. The behavior of an animal can be shaped when specific actions are paired with reinforcing stimuli, and operant learning occurs through the selection of actions that result in rewarding consequences. Although there is some evidence of the neurophysiological components underlying this selection mechanisms, it is still not fully understood how the process of selection occurs within neuronal circuits. In the previous chapter we have discussed how specific features of behavior can be selected based on reinforcement, by selecting specific neuronal patterns that produce outcome-relevant behaviors. In this study we aimed at testing if selection could be introduced directly at the neuronal circuits level, by pairing a specific pattern of activity with dopaminergic activation, in order to shape neuronal and behavioral activity. Here we

describe a closed-loop brain-machine-interface paradigm for mice, in which pairing a specific pattern of cortical activity with optogenetic activation of midbrain dopaminergic cells, gives rise to a bias in the neuronal activity towards the reinforced pattern. These data could provide some insight into the neurophysiological mechanisms that underly the selection of particular neural patterns that lead to relevant outcomes.

INTRODUCTION

In 1898, Edward Thorndike put forward his “law of effect”, which states that any behavior followed by pleasant consequences is more likely to be selected and repeated, while behaviors that are followed by unpleasant consequences are more likely to be suppressed (Thorndike, 1898). Based on Thorndike’s work, Skinner (Skinner, 1938) coined the term operant conditioning, demonstrating that we can shape the behavior of animals by the use of reinforcements following the desired behavioral response.

Learning to perform a motor skill, from riding a bicycle to playing the piano, usually comprises a process of trial and error. Through this process, an animal initially explores the behavioral space in which it is inserted, and based on the consequences of this exploration, selects the motor patterns that lead to the desired output (Costa, 2011). Nevertheless, given that there is no mechanism to directly reinforce particular muscular patterns, it is proposed that behavioral selection might be caused by a similar process of neuronal pattern selection.

We and other groups have previously shown that neuronal dynamics change during skill learning, with the variability of striatal and cortical circuits varying in a way that is consistent with the theory of exploration

and selection for motor skill learning. Work in song-birds has demonstrated that the decrease in song variability is accompanied by a decrease in neuronal variability until a crystalized pattern is established (Hahnloser et al., 2002). Variability in corticostriatal circuits also decreases from early to late learning, as mice learn and consolidate a motor skill (Costa et al., 2004), and in primates, learning to control a neuronal prosthetic device leads to the emergence of a crystalized neuronal representation (Ganguly and Carmena, 2009). Task-relevant behavior dimensions are specifically optimized during motor learning and, as described in the previous chapter, there is evidence that neuronal activity dynamics might follow closely the refinement of outcome-relevant re-organizations. These observations suggest that, as we learn and automatize novel actions, there is a mechanism of selection of neuronal patterns that reinforces activity leading to the desired outcomes.

Dopamine is known to be critical for the generation of novel actions and for the modulation of variability within corticostriatal circuits. The phasic activity of midbrain dopamine neurons provides a reward prediction error, used to guide learning, through projections to the cortex and basal ganglia. Classical studies of intracranial electrical self-stimulation using operant tasks (Olds, 1958; Wise, 1981) have shown that activation of dopaminergic cells can substitute natural reinforcements and promote the selection of specific behaviors. More recent studies using selective optogenetic stimulation of the dopaminergic pathways have promoted further insight into these mechanisms by showing that selective activation of both VTA (Adamantidis et al., 2011; Tsai et al., 2009) or SNc (Rossi et al., 2013) dopaminergic neurons, in substitution of a natural reward, is sufficient to promote instrumental learning.

Nevertheless, the mechanism of selection within neuronal circuits during motor learning still lacks strong evidence and support. Brain-machine-interface paradigms, by circumventing the normal biological pathways involved in motor control, allow to directly test the volitional modulation and control of neuronal signals that normally have a complex relationship to behavior or cognition (Fetz, 2007), and have recently been shown to be a reliable tool to promote insight into the mechanisms of motor skill learning (Orsborn and Carmena, 2013).

The present study aims at evaluating if the process of selection can be observed directly within the neuronal networks relevant to a task, using a closed-loop brain-machine-interface paradigm.

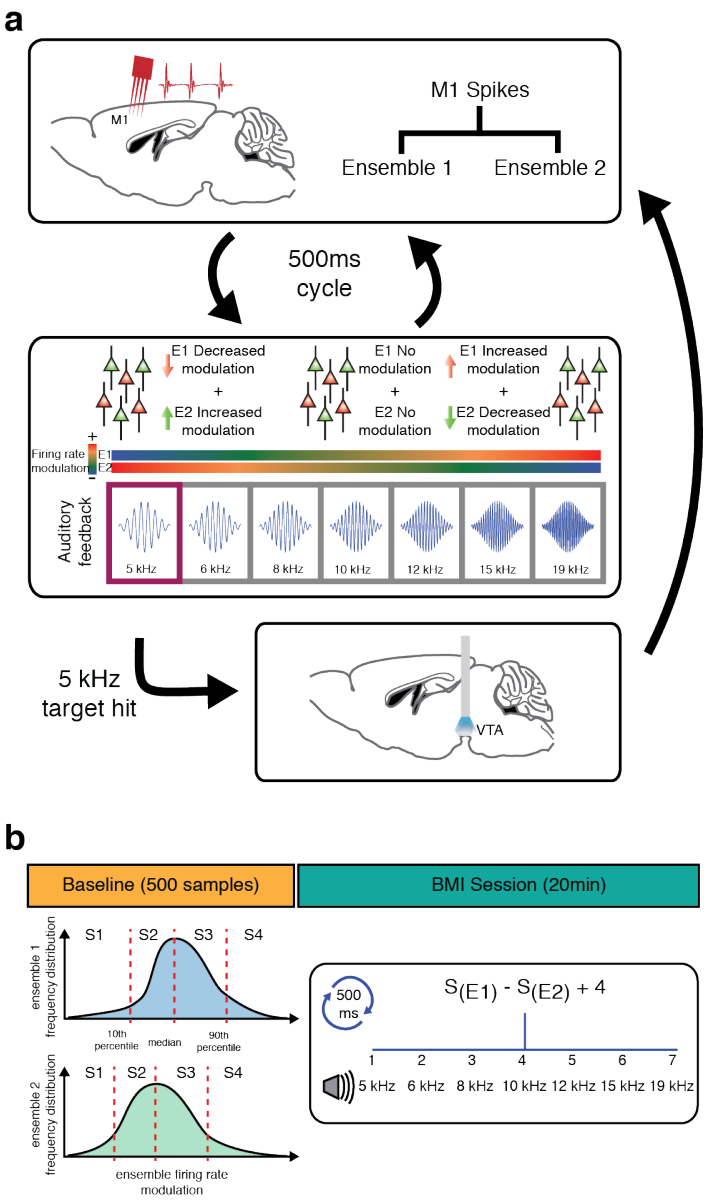


Figure 3.1 | Brain-machine-interface paradigm design. (**a**) Diagram of the online BMI task. Neuronal spike timestamps from selected M1 units were attributed to one of two ensembles that were used as input for an online algorithm. Each 500ms the modulation of the two ensembles was evaluated and translated into one of seven audio frequencies that were played back to the animals. Reaching the lower frequency target (5kHz) would trigger a blue laser stimulation of the dopaminergic cells of the VTA. (**b**) Daily baseline calculation. Prior to each training session, we recorded 500cycles to evaluate the baseline modulation for each ensemble, defining 4 modulation states per ensemble, which allowed the online estimation of firing rate and translation into the appropriate auditory frequency

RESULTS

We have developed a brain-machine-interface paradigm for mice based on the work of Koralek and colleagues (Koralek et al., 2012), where the activity of two ensembles of 2-4 units from the primary motor cortex (M1), that were arbitrarily defined as outcome-relevant, was translated by an online algorithm into an auditory frequency (**Fig. 3.1a**). Mice had to precisely control the modulation of the outcome-relevant cells in order to reach a specific pattern of neuronal activity within these ensembles. Online estimations of firing rate modulation were mapped each 500ms into one of seven different auditory frequencies, which were played back to the mice. By decreasing firing rate modulation of the first ensemble, while concomitantly increasing the modulation in the second ensemble, mice would drive the audio feedback towards low frequencies. On the contrary, increasing firing rate modulation of the first ensemble while decreasing the modulation of the second ensemble, would drive the feedback towards higher frequencies.

The lower pitch frequency (5kHz) was defined as the target, and it was paired with a train of 10ms light pulses (at 14Hz for 2 seconds) from a blue laser, resulting in the stimulation of dopaminergic cells in the ventral tegmental area contralateral to the recording site. When none of the two boundary frequencies (5 or 19kHz) were reached within 60 seconds, mice were given a timeout of 10 seconds with no audio feedback or stimulation.

Mice were infected with a virus that promoted the expression of channelrhodopsin-2 in the dopaminergic cells of the VTA contralateral to the recording site (ChR2 group, N=10), and underwent normal BMI training, with dopaminergic stimulation substituting for a natural reward after target reach. Control mice were infected with a viral vector that promoted the expression of a fluorescent marker within the

dopaminergic cells of the VTA (YFP group, N=7), but no light sensitive channel, undergoing normal BMI training but being insensitive to the light stimulation.

After 3 days of BMI training (one daily session with 20min duration), a significant increase in the percentage of hits for the target frequency, compared to the opposite boundary frequency (19kHz), became evident for the ChR2 group (significant interaction between training session and frequency, $F_{3,54}$=2.838, p=0.0465) but not for the YFP control group (no significant interaction between training session and frequency, $F_{3,36}$=1.071, p=0.3737) (**Fig. 3.2a**). These relative change values were calculated for each individual animal, relatively to the percentage of boundary frequencies hits of the first training session.

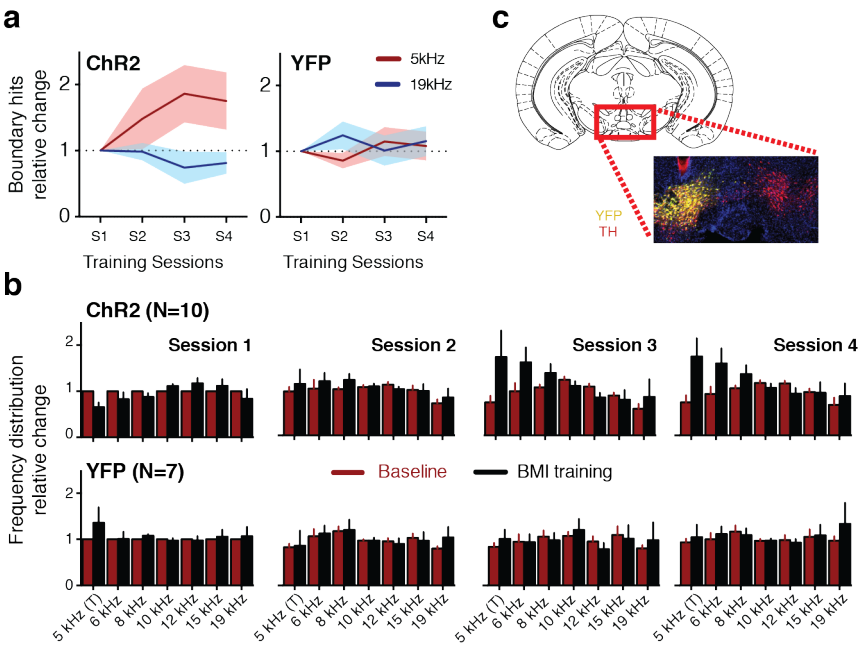


Figure 3.2 | Reinforcement of a specific pattern of neuronal activity by dopaminergic stimulation promotes a shift in the auditory feedback. (**a**) Relative change in the percentage of hits, for both boundary frequencies (5kHz and 19kHz) and for both groups (ChR2 and YFP). Color shades represent mean ± SEM (**b**) Histograms representing the relative change in the frequency distribution for the 7 possible auditory frequencies. Top row corresponds to 4 sessions of the ChR2 group and bottom row to the sessions of the YFP group. Data for both the baseline and BMI training of each day is represented. Error bars represent mean ± SEM (**c**) Coronal brain slice depicting the viral infection specific to the dopaminergic cells of the VTA, with the immunohistochemistry labelling for both tyrosine hydroxylase (TH, red) and the cre-dependent fluorescent protein (YFP, yellow).

Further analysis of the impact of selective dopaminergic stimulation on the output of our BMI task, clearly demonstrates that after as few as 3 days of training there is a shift in the distribution of the output states towards the target frequency, which was paired with the dopaminergic stimulation (**Fig. 3.2b**). Comparison of the output of BMI training sessions, with the output obtained by running our algorithm with neuronal data recorded during the baseline periods, shows a significance difference between the two blocks (ChR2 group, session 3; significant main effect for baseline vs training: $F_{1,63}$=4.726, p=0.0335),

with an increase in percentage of target frequency hits, during BMI training compared to baseline periods (Bonferroni's multiple comparisons test: t_{63}=3.279, p=0.0119, **Fig. 3.2b**). YFP controls did not exhibit any difference between baseline and BMI training periods (YFP group, session 3; no significant main effect for baseline vs training: $F_{1,42}$=0.07422, p=0.7866, **Fig. 3.2b**).

These data suggest that indeed, pairing a specific neuronal pattern with the phasic activation of midbrain dopaminergic neurons is sufficient to promote selection of neuronal patterns, in a mechanism that might underlie behavioral selection during motor skill learning.

DISCUSSION

Behavioral selection refers to the process that allows animals to select from the multitude of actions that results from motor exploration. This interplay between motor exploration and action selection is known to be vital for learning and acquisition of novel skills, and to rely on dopamine dependent processes. Artificial activation of dopaminergic cells, either electrically (Olds, 1958; Wise, 1981) or using light (Adamantidis et al., 2011; Rossi et al., 2013; Tsai et al., 2009), is sufficient to reinforce behavior, promoting the selection of specific behavioral patterns. Work in songbirds suggests that dopamine can provide a reinforcement signal to the song performance, shaping neuronal and behavioral patterns (Goldberg and Fee, 2011), and in humans, the selection of low-level movement parameters has been shown to be under the influence of dopaminergic processes (Galea et al., 2013; Mazzoni et al., 2007).

In parallel with the behavioral dynamics observed throughout the process of motor learning, several studies have provided some insight into the neurophysiological hallmarks of exploration and selection, with

observations of high levels of neuronal variability during early learning phases and subsequent decrease in neuronal variability accompanying the selection of appropriate motor patterns. In humans, increased behavioral variability has been shown to correlate with corticospinal excitability, in a process dependent on dopamine (Galea et al., 2013). During motor skill learning, mice decrease the variability within corticostriatal circuits (Costa et al., 2004), with dopamine also playing a critical role for shaping and selecting the correct neuronal and action patterns (Costa, 2011; Costa et al., 2006). A study by Mazzoni and colleagues (Mazzoni et al., 2007) gives further support to the importance of dopamine in the process of action selection, by showing that patients with Parkinson's disease implicitly select slower movements, even while maintaining the ability to perform accurate fast movements.

In this study we aimed at testing whether we could artificially induce selection of a specific neuronal pattern, in ways similar to those that operate under normal motor learning and during action selection. We demonstrate that using a brain-machine-interface closed-loop paradigm to directly reinforce an arbitrarily defined outcome-relevant neuronal pattern, through activation of midbrain dopaminergic neurons in the Ventral Tegmental Area (VTA), we can bias the neuronal activity patterns towards the reinforced target. Within a few training sessions, closed-loop reinforcement of a specific pattern of motor cortex activity by dopaminergic stimulation, promoted a shift in the auditory feedback that translated the neuronal activity modulations.

These data suggest, that during learning, a very precise temporal or spatio-temporal pattern can emerge and be selected by error signals. This is in agreement with observations from songbirds were crystalized patterns of activity are selected as the animals learn and practice a song

(Kao et al., 2005; Olveczky et al., 2005), and from primates learning to control neural prosthetics, were stable maps of cortical activity emerge as they learn the task (Ganguly and Carmena, 2009).

This provides support to the hypothesis that selection of outcome-relevant neuronal patterns, through dopaminergic-dependent processes, might be the basis underlying the behavioral variability dynamics observed during motor skill learning.

METHODS

Animals

All experimental procedures were carried in accordance to the ethics committee guidelines of the Champalimaud Foundation and with approval of the Portuguese DGAV. Experimental procedures were carried out using seventeen male, 3 to 5 month, BAC transgenic mice expressing Cre recombinase under the control of tyrosine hydroxylase (FI12) (Gong et al., 2007). Animals were maintained on a light-dark cycle of 12h:12h starting at 7AM. All experiments were done during the light cycle. Mice were housed in groups of 4 animals prior to surgery and individually after electrode implantation and viral injection.

Surgical implantation

Seventeen animals were unilaterally implanted with 16 tungsten electrodes (ø 23-35μm) with micro-polished tips. The electrodes were arranged either fixed in a 2x8 array or a movable bundle of 16 wires. Craniotomies and electrode array positioning were done according to coordinates from the Mouse Brain Atlas (Paxinos and Franklin, 2004). The array was targeting the primary motor cortex (M1, layer 5) and placed under the following coordinates: 1mm rostral and 1.6mm lateral

from bregma, and lowered ~ 1mm from the surface of the brain. Electrodes were manually lowered at slow rates while constantly monitoring neural activity in all the channels in order to control for proper electrode function and correct positioning. Final verification of electrode position was done after all the experimental procedures were finished, by perfusing animals with PFA and histological confirmation of Nissl stained 70µm brain slices.

For viral expression of ChR2 in Th-cre mice, 1µl of a cre-inducible adeno-associated virus (AAV) vector carrying the gene encoding the light-activated cation channel channelrhodopsin-2 and a yellow fluorescent reporter (AAV2/1.EF1a.DIO.hChR2-eYFP, titer $4.04x10^{12}$) was stereotactic delivered into the ventral tegmental area (VTA) contralateral to the recording electrode, enabling specific expression of ChR2 in midbrain dopaminergic cells (N=10 mice) (**Fig. 3.2c**). Control animals (N=7 mice) were injected with a viral vector that only carried the yellow fluorescent reporter (AAV2/1.EF1a.eYFP, titer $1.85x10^{12}$). The following stereotaxic coordinates were used to target the VTA: 3.16 mm caudal and 0.48 mm lateral from bregma and lowered 4.4 mm from the surface of the brain. Viral injection was done through a glass pipette using a syringe pump (Nanoject II, Warner Instruments) precisely controlled by 10 ms electrical pulses at 0.2 Hz, where each pulse triggered a 4.6 nl solution injection. The pipette was left in position for ~20 min after the injection and then slowly moved out. Subsequently, a 200 mm diameter fiber optic (NA 0.22) was implanted 0.2 mm above the viral injection. Following the implantation, a plastic cap was used to cover the fiber, and mice were placed in the home cage for 15 days, allowing both viral expression and surgery recovery, before further experiments.

Electrophysiology

Single and multi unit activity was recorded using Blackrock Microsystems Neural Signal Processor (NSP), allowing for online sorting of identified units. Neuronal activity was sorted using the online sorting application prior to each daily recording session. Only units with a clearly identified waveform and high signal-to-noise ratio were used. Further offline sorting of selected units was done using Plexon Offline Sorter v3 (Plexon, Inc), based on waveform characteristics, inter-spike intervals and PCA clustering. Timestamps from the operant box were input into the NSP and synchronized with the neuronal data. The NSP was connected and controlled by a MATLAB (The MathWorks, Inc) custom program using the CBMEX library (Blackrock Microsystems).

Brain-machine-interface paradigm

Animals were exposed to the BMI paradigm training two weeks after surgery. Behavioral sessions took place in an operant box with controllable lights (Med Associates, Inc). Recorded neural data was sent in real time to a custom program in Matlab, which converted the neural modulations into appropriate auditory feedback, played through a speaker mounted on the side of the operant box. Frequencies used for auditory feedback ranged from 5 to 19 kHz in quarter-octave increments, matching rodent discrimination thresholds (Han et al., 2007).

Activity of the recorded units was binned in 500 ms bins and 2 ensembles of 2-4 cells were defined based on waveform, ISI histogram and signal-to-noise ratio, to be used as input for the BMI controller. An online algorithm translated neural modulations of the selected ensembles into the pitch of an auditory cursor each 500ms, and by modulating activity in these two ensembles, mice controlled the auditory feedback frequency (**Fig. 3.1a**). Baseline activity, based on 500 cycles were no auditory feedback was given, was determined each day prior to

BMI training, in order to assess the distribution of firing rate modulation for each ensemble. Activity of individual units was z-scored by subtracting the median firing rate and dividing it by the range. Ensemble modulation was defined by the sum of the z-scored firing for each individual unit that composes the ensemble. Based on this baseline, 4 modulation states (S1-S4) were defined for each ensemble, using the median ensemble modulation and the 10^{th} and 90^{th} percentiles. During the task, the combined modulation of the two ensembles (S_{e1}-S_{e2}+4) was calculated for each 500ms cycle and translated into the appropriate pitch for the auditory feedback (steps 1-7 ~ 5-19kHz) (**Fig. 3.1b**). To create some smoothness on the auditory feedback we performed a weighted average of 3 bins at each cycle. The same units were kept in each ensemble across training days, unless signal quality decreased considerably, in which case a neighbour unit with better signal quality would substitute individual units.

When the target frequency (5kHz) was hit, a Data Acquisition I/O board (National Instruments, Inc) would trigger the delivery of blue light stimulation through the optic fiber, using a 473nm laser. The stimulation train consisted of 10ms duration pulses delivered at 14Hz for 2 seconds. The laser power used for stimulation of dopaminergic cells was ~10mW measured at the tip of the optical fiber.

Data Analysis and Statistics

All analyses and statistics were done in MATLAB (The Mathworks, Inc) using custom written programs, or using GraphPad Prism (GraphPad Software, Inc).

To evaluate the relative change in boundary hits, we have compared the percentage of cycles that fell within each of the two boundaries (5 and 19kHz) out of all the cycles that reached the extreme frequencies, and subsequently normalized it, using the data of the first session as

reference.

To determine the frequency distribution of the BMI controller, we first calculated for each animal and session the percentage of cycles that fell within each of the 7 output states, and quantified the relative change in these values across sessions, using data from the first training session as our reference. In one YFP animal, where the percentage of cycles reaching the target frequency was considerably low (0.5%) for the first session, data from the first baseline was used as reference for subsequent sessions. The relative change in the first session for all animals was calculated using the baseline data of that same session as reference.

Two-way ANOVA was used for evaluating differences between the relative change of the boundary hits, and differences between relative change of auditory frequencies in session and baseline periods.

ACKNOWLEDGMENTS

We thank J. Carmena, A. Koralek and V. Athalye for help and support setting up the BMI paradigm, and A. Vaz for mouse colony management. This research was supported by the INDP Graduate Programme and a FCT fellowship to F.J.S., and an European Research Council Starting Grant and a HHM1 International Early Career Scientist Grant to R.M.C.

AUTHOR CONTRIBUTIONS

F.J.S. performed all the experimental work and analyses. F.J.S. and R.M.C. designed the experiments and wrote the manuscript.

REFERENCES

Adamantidis, A., Tsai, H.C., Boutrel, B., Zhang, F., Stuber, G.D., Budygin, E.A., Tourino, C., Bonci, A., Deisseroth, K., and de Lecea, L. (2011). Optogenetic Interrogation of Dopaminergic Modulation of the Multiple Phases of Reward-Seeking Behavior. J. Neurosci. *31*, 10829–10835.

Costa, R.M. (2011). A selectionist account of de novo action learning. Curr. Opin. Neurobiol. *21*, 579–586.

Costa, R.M., Cohen, D., and Nicolelis, M.A.L. (2004). Differential corticostriatal plasticity during fast and slow motor skill learning in mice. Curr. Biol. *14*, 1124–1134.

Costa, R.M., Lin, S.-C., Sotnikova, T.D., Cyr, M., Gainetdinov, R.R., Caron, M.G., and Nicolelis, M.A.L. (2006). Rapid alterations in corticostriatal ensemble coordination during acute dopamine-dependent motor dysfunction. Neuron *52*, 359–369.

Fetz, E.E. (2007). Volitional control of neural activity: implications for brain-computer interfaces. J Physiol (Lond) *579*, 571–579.

Galea, J.M., Ruge, D., Buijink, A., Bestmann, S., and Rothwell, J.C. (2013). Punishment-Induced Behavioral and Neurophysiological Variability Reveals Dopamine-Dependent Selection of Kinematic Movement Parameters. J. Neurosci. *33*, 3981–3988.

Ganguly, K., and Carmena, J.M. (2009). Emergence of a stable cortical map for neuroprosthetic control. PLoS Biol. *7*, e1000153.

Goldberg, J.H., and Fee, M.S. (2011). Vocal babbling in songbirds requires the basal ganglia-recipient motor thalamus but not the basal ganglia. J. Neurophysiol. *105*, 2729–2739.

Gong, S., Doughty, M., Harbaugh, C.R., Cummins, A., Hatten, M.E., Heintz, N., and Gerfen, C.R. (2007). Targeting Cre recombinase to specific neuron populations with bacterial artificial chromosome constructs. J. Neurosci. *27*, 9817–9823.

Hahnloser, R.H.R., Kozhevnikov, A.A., and Fee, M.S. (2002). An ultra-sparse code underlies the generation of neural sequences in a songbird. Nature *419*, 65–70.

Han, Y.K., Köver, H., Insanally, M.N., Semerdjian, J.H., and Bao, S. (2007). Early experience impairs perceptual discrimination. Nat. Neurosci. *10*, 1191–1197.

Kao, M.H., Doupe, A.J., and Brainard, M.S. (2005). Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. Nature *433*, 638–643.

Koralek, A.C., Jin, X., Long, J.D., Costa, R.M., and Carmena, J.M. (2012). Corticostriatal plasticity is necessary for learning intentional neuroprosthetic skills. Nature *483*, 331–335.

Mazzoni, P., Hristova, A., and Krakauer, J.W. (2007). Why don"t we move faster? Parkinson"s disease, movement vigor, and implicit motivation. J. Neurosci. *27*, 7105–7116.

Olds, J. (1958). Self-Stimulation of the Brain Its Use To Study Local Effects of Hunger, Sex, and Drugs. Science *127*, 315–324.

Olveczky, B.P., Andalman, A.S., and Fee, M.S. (2005). Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. PLoS Biol. *3*, e153.

Orsborn, A.L., and Carmena, J.M. (2013). Creating new functional circuits for action via brain-machine interfaces. Front Comput Neurosci *7*, 157.

Paxinos, G., and Franklin, K.B.J. (2004). The Mouse Brain in Stereotaxic Coordinates (Gulf Professional Publishing).

Rossi, M.A., Sukharnikova, T., Hayrapetyan, V.Y., Yang, L., and Yin, H.H. (2013). Operant Self-Stimulation of Dopamine Neurons in the Substantia Nigra. PLoS ONE *8*, e65799.

Skinner, B.F. (1938). The Behavior of Organisms (Oxford: Appleton-Century).

Thorndike, E.L. (1898). Animal intelligence: An experimental study of the associative processes in

animals. The Psychological Review: Monograph Supplements *2*, i–109.

Tsai, H.-C., Zhang, F., Adamantidis, A., Stuber, G.D., Bonci, A., de Lecea, L., and Deisseroth, K. (2009). Phasic Firing in Dopaminergic Neurons Is Sufficient for Behavioral Conditioning. Science *324*, 1080–1084.

Wise, R.A. (1981). Intracranial self-stimulation: mapping against the lateral boundaries of the dopaminergic cells of the substantia nigra. Brain Res. *213*, 190–194.

4

DISCUSSION

The goal of this dissertation was to provide further insight into the mechanisms of motor skill learning and to provide a new framework for the experimental data that has been collected throughout the past years. The findings within this dissertation advance our understanding of the role of neuronal variability and reinforcement in motor learning.

Action generation and performance encompass 3 main features: motivation (why we do the action), policy / approach (what we do) and control (how to do the action). Frontal cortical areas and the limbic system are important players in the process of decision making and the motivation for action (Gerbner, 1972; Kornhuber, 1978; Mogenson et al., 1980; Seo et al., 2012). The basal ganglia and the cerebellum are involved in selecting which motor program and strategies to use for the task at hand (Hikosaka, 1998; Kropotov and Etlinger, 1999; Marsden, 1984). And finally, the cortex, brainstem, and spinal cord are primarily involved in how to perform the action and the execution of the movements (Favorov et al., 1988; Fetz et al., 2002; Georgopoulos, 1996). Our work and this dissertation have focused mainly on the processes occurring during the selection and execution phases, and we provide novel behavioral and neuronal evidence, together with the studies from Jin and Costa (Jin and Costa, 2010), for the involvement of the striatum in the process of how to perform motor actions.

In Chapter 2, we have demonstrated that learning a motor skill has several behavioral and neuronal hallmarks and that the interplay

between these two levels can follow some principles that are aligned with recent theories of motor control. By splitting the characterization of the variability in our motor task into outcome-relevant and outcome-irrelevant dimensions we provide a more systematic way of looking into behavioral and neuronal correlates. We have observed differential modulation of outcome-relevant and irrelevant features, with behavioral refinement strategies occurring in accordance with theories of optimal feedback as a framework for motor control. After describing the differential behavior re-organization, we recorded activity in corticostriatal circuits during the process of learning. Cortico-basal ganglia circuits and loops are known to be involved in different stages of the motor learning process (Costa et al., 2004; Karni et al., 1998; Luft and Buitrago, 2005; Ungerleider et al., 2002). Several studies have shown differential recruitment of specific regions across the learning process, as well as a decrease in variability of these circuits as animals learn and acquire different sets of skills (Costa et al., 2004; Kao et al., 2005). We have shown that the observed reduction in neuronal variability is correlated with the dynamics for the outcome-relevant behavioral features and not with the outcome-irrelevant features. This points to the hypothesis that the corticostriatal circuits can encode information relative to the selection of relevant features for a task, optimizing behavior and neuronal activity related to those specific features. In Chapter 3, we expand this further by demonstrating that closed-loop reinforcement through stimulation of VTA populations of cells is sufficient to promote selection of a specific pattern of neuronal activity.

The ability that humans and other animals have to generate and perform movements and actions, in ever-evolving and challenging environments, requires an amazing capacity to constantly acquire and learn novel and

complex motor skills. There are different views of how animals learn to generate particular actions in particular situations: instructionist views defend that information from the environment is transferred into the brain creating completely novel neural and behavioral patterns, while selectionist views propose differential amplification of existing ongoing neural and behavioral patterns that better fit the situation (Sporns and Tononi, 1994). Edelman (Edelman, 1987) has used these concepts to formulate the theory of Neuronal Darwinism by establishing parallelisms with the immune system and evolution. His theory postulates that the brain is mainly a selectionist system instead of an instructional one, similarly to what happens in evolution, but on a somatic timescale. These neuronal group selection mechanisms occur in two major stages: a developmental selection stage, during which adjacent cells become interconnected and established into neuronal circuits; and a second stage of experimental selection, during which the activity of the animal and the interaction with the external world promotes the adjustment of the synaptic connections strengths that yield a most efficient behavior (Edelman, 1987). These theories predict that under situations where an animals needs to generate novel actions, the nervous system would generate novel patterns of activity, resulting in variable and novel actions that can then be selected during the process of motor learning, to better meet task requirements (Changeux, 1989; Edelman, 1987). A study by Dragoi and Tonegawa (Dragoi and Tonegawa, 2010) provides strong evidence to the selectionist theory, by demonstrating that the patterns selected during learning can already be determined by the neuronal activity during resting periods.

In this dissertation we demonstrate that during motor skill learning variability of outcome-relevant features is increased in early phases, and is then selected and refined as learning progresses. Furthermore, neuronal variability within corticostriatal circuits is also selected during

learning, as it increases in early sessions and decreases as training progresses. These neuronal variability dynamics correlate specifically with the outcome-relevant behavioral dimensions. Additionally, we show that the selection of cortical activity patterns can occur by direct stimulation of midbrain dopaminergic cells.

A common trait in operant learning and reinforcement learning is the critical balance between exploration and exploitation (Sutton and Barto, 1998). Exploration consists of discovering new features about the task by choosing sub-optimal actions, while exploitation is the act of selecting the best action known so far, and thereby taking advantage of knowledge gained during previous attempts. Early learning phases are usually described as more exploratory, hence with higher behavioral variability, as the animal explores the world around him. Animals gradually shift towards more exploitive strategies in the latter learning phases, selecting behaviors that increase the likelihood of success, hence decreasing the variability in the behavioral output.

The cortico-basal ganglia circuits comprise a set of neuronal loops that based on their anatomical organization and functional features provide a suitable candidate for the generation of the variability that underlies motor and behavioral variability. The striatum, the entry point of the basal ganglia, receives a wide range of inputs from nearly all the cerebral cortices. The associative and sensorimotor cortices have the most prominent contributions, with projections displaying a topographic distribution throughout different areas of the striatum. The putamen, or DLS in rodents, receives mainly inputs from sensorimotor areas like the premotor motor and sensorimotor cortex; the caudate, or DMS in rodents, receives most of its input from the associative areas, like the prefrontal, temporal, parietal and cingulate cortex, and the motor areas

controlling eye movements (Haber, 2003; McGeorge and Faull, 1989; Voorn et al., 2004). Inputs from the sensorimotor and motor cortices tend to converge and overlap within the ipsilateral striatum, intercalating with inputs from the contralateral motor areas (Flaherty and Graybiel, 1991; 1993). The primary motor cortex in turn, receives a major projection from the ventrolateral nucleus of the thalamus (VL) (Asanuma et al., 1974; Hunnicutt et al., 2014). This projection is diffuse, with one VL neuron projecting to a wide area of the motor cortex, hence not being a likely candidate for the organized contraction of muscles seen in skill learning (Asanuma and Keller, 1991). The projection from sensory to motor cortex on the other hand is highly specific (Porter and Sakamoto, 1988). Evidence from animal models demonstrates that removal of sensorimotor cortex impairs acquisition of motor skill, but its removal after training results in no change in performance (Pavlides et al., 1993). Specific lesions of the primary motor cortex also cause severe motor skill deficits (Friel and Nudo, 1998; Hoffman and Strick, 1995).
These anatomical and lesion experiments demonstrate the importance of the cortico-basal ganglia circuits in the process of motor skill learning.

The basal ganglia circuits can exhibit spontaneous activity (Aldridge et al., 1990; Elias et al., 2008; Plenz and Kital, 1999) and are susceptible to several types of modulation, which can promote the generation of neuronal variability. Such variability is accepted to be critical for learning novel motor skills and to underlie the behavioral variability associated with the process of trial and error during motor skill learning. Indeed, studies in several animal models have demonstrated that early learning phases have increased variability (Costa et al., 2004; Kao et al., 2005; Mandelblat-Cerf et al., 2009), and a more recent study in humans provided evidence that higher baseline variability levels can promote faster learning in a motor task (Wu et al., 2014). This is consistent with

the idea of increased motor exploration as a fundamental feature of procedural learning.

Learning complex motor skills is often designated as sequence learning as movements become organized in sequences or chunks (Asanuma and Pavlides, 1997). Sequence learning is comprised of two separate features: acquisition of the order of elements that comprise the sequence, and the increase in performance for each of the single elements that belong to the sequence (Kitago and Krakauer, 2013). The involvement of the basal ganglia in learning and consolidating action sequences has been well established (Brainard and Doupe, 2002; Graybiel, 1998; Hikosaka et al., 1998), but there is still some controversy on the exact mechanisms and the roles that these areas have. One aspect that contributes to the disparity of roles attributed to the basal ganglia in learning and control of movements comes from the fact that many studies use motor behavior as a readout of higher-order features of behavior (e.g. (Hikosaka et al., 1995)). Several studies have investigated the role of the striatum on sequence-learning through the impact of striatal lesions and inactivations (Eckart et al., 2010; Hikosaka et al., 1995; Miyachi et al., 1997; Moussa et al., 2011). Nevertheless, since in most studies the required movements are easy and do not require any speed-accuracy tradeoff, they do not provide a clear distinction between the effects on high-order features (e.g. sequence order and structure) and the effects on the quality of the movements proper (e.g. (Eckart et al., 2010; Hikosaka et al., 1995; Miyachi et al., 1997).

Studies in songbirds have also provided much evidence for the role of the basal ganglia in sequence learning. Song learning in birds resembles in several ways some of the mechanisms observed in human motor

learning (Thorpe, 1958), specially speech learning (Brainard and Doupe, 2002; Doupe and Kuhl, 1999; Marler, 1970). Young birds learn to produce their adult song in two phases: a sensory phase, in which they listen to the song of a tutor bird; and a sensorimotor phase, in which young birds try to replicate the memorized template song and refine their own song based on auditory feedback. The initial vocalizations (also called "subsong") are highly variable and unstructured, and as the animal practices the song, the subsong becomes more structured, less variable and often indistinguishable from the template song, in a process known as "crystallization" (Fee and Scharff, 2010; Marler, 1981; Williams, 2004). The brain areas associated with learning and performance of the song have been extensively characterized. Both Area X (the avian homolog of basal ganglia) and its cortical target (lateral magnocellular nucleus of the anterior neostriatum, LMAN) have been implicated in song learning. Disruption of these areas impairs learning in juvenile birds, but has no effect on the performance of adults. Lesions of LMAN promote a loss of variability and premature crystallization, resulting in a poor imitation of the tutor song (Bottjer et al., 1984; Scharff and Nottebohm, 1991; Sohrabji et al., 1990), lesions of Area X also lead to a poor imitation of the tutor song, but in contrast to the effect in LMAN, the produced song maintains a high level of syllable variability and does not crystallize with practice (Scharff and Nottebohm, 1991). Accurate feedback of the motor output is critical for birds to properly perform their songs, and to adjust specific features of their song to avoid disruptive feedback (Sakata and Brainard, 2006). Variability in song production is critical for this adaptation and optimization of relevant behaviors (Tumer and Brainard, 2007).

In Chapter 2, we have presented a task in which mice were required to perform a complex sequential action. We have removed the effect of sequence order in this task by using a homogenous sequence, which

reduces the number of behavioral features controlled by the animal. We show that variability of outcome-relevant features is reduced with training, while no reduction in variability of outcome-irrelevant dimensions was observed. Cortical and striatal variability during sequence performance vary throughout learning and correlate specifically with the dynamics observed for the behavioral outcome-relevant dimension. Our data, together with the evidence from songbird studies, provide further support to the generation and modulation of variability as critical features for learning of motor skills.

The activity of the cortico-basal ganglia circuits can be modulated by multiple factors, most prominent of which are dopaminergics inputs (Bolam et al., 2000; Gerfen and Surmeier, 2011; Haber, 2014; Surmeier et al., 2011). There are two main dopaminergic nuclei in the midbrain: the ventral tegmental area (VTA) and the substantia nigra pars compacta (SNc). These two populations of dopaminergic cells are believed to have different functions: VTA cells have been demonstrated to carry error signals to the basal ganglia as a major component of reinforcement learning (Schultz et al., 1997) and a critical part of the process of selection; in turn, SNc populations in turn have been implicated in the generation of variability and novel patterns in the cortico-basal ganglia circuits (Jáidar et al., 2010). The death of SNc neurons is one of the hallmarks of Parkinson's disease, in which patients show an impairment of the ability to generate and initiate novel voluntary actions (Carlsson, 1972). Depletion of dopamine in mice leads to the absence of self-paced movements; on the contrary, high dopaminergic levels promote the appearance of hyperkinetic states (Costa et al., 2006). Therefore dopamine is accepted as a critical modulator of the activity within the cortico-basal ganglia circuits and has a major impact on the process of motor learning and action generation.

The view that learning occurs through trial and error and that neuronal variability underlies the generation of novel neuronal patterns, which allow for the exploration of the task space, leads to the prediction that the dynamics of neuronal variability change through learning. In fact, as we have demonstrated in Chapter 2 and supported by other studies, variability in the cortico-striatal circuits is increased during early phases of learning (Barnes et al., 2005; Costa et al., 2004). As the subjects acquire the motor tasks and consolidate the motor memories, variability in these circuits decreases, accompanying the increases in performance. In addition to this, with skill learning there are changes in the modulation of these circuits by dopamine, which is suggested to be a variability generator. As training progresses, LTP in D2 receptor-expressing cells increases, and the performance of the skill becomes less dependent on the activation of D1 receptors. (Yin et al., 2009). In Chapter 2 we put forward the hypothesis that the dynamics observed in the variability of the neuronal circuits underlie the refinement of outcome-relevant dimensions. This might help explain also some of the divergent functions attributed to the basal ganglia and the cortico-striatal loops. In the task we have presented in Chapter 2, the frequency of a sequence of lever-presses was the outcome-relevant feature, and this was the only behavioral feature that was optimized by the animals, in opposition to the non-relevant features where variability increased. Furthermore, the neuronal dynamics observed within the cortico-striatal circuits correlated specifically with the variability of frequency of the sequences of actions. Many studies have found that the basal ganglia is involved in several other functions including eye movements, cognitive functions and emotional processes (Arsalidou et al., 2013; Kotz et al., 2009; Marchand, 2010; Shires et al., 2010; Whishaw et al., 1987). The differences in these observations could potentially be explained by the

different tasks used in all these experiments. The striatum and the corticostriatal circuits have been proposed as critical for action selection and the decision of which action to perform from all the possible behaviors for a specific context (Costa, 2011; Grillner et al., 2013). We demonstrate in Chapter 2, that plasticity in the corticostriatal circuits is crucial for selection of the outcome-relevant features of our task. Here we put forward the hypothesis that this could be a general mechanism, such that the basal ganglia and the cortico-basal ganglia loops promote the selection of the relevant feature for each task. The fact that lesions within the basal ganglia exhibit such strong motor deficits led to the acceptance that the major function of the basal ganglia might be related to general motor execution. However, when tested in tasks where the task-relevant feature is non-motor (Kotz et al., 2009), basal ganglia signals relate to the specific aims of each task. Taken together the work presented in this dissertation supports the emerging view that the basal ganglia are capable of remarkable behavioral specificity.

Motor learning and performance require a complex interaction between the mechanical proprieties of the body and the circuits composing the nervous system. In 2002, Todorov and Jordan, in an attempt to understand how multiple biomechanical degrees of freedom are coordinated for proper motor control and execution, proposed a theory based on optimal feedback theory as a mechanism for motor control (Todorov and Jordan, 2002). This framework establishes several predictions for the process of motor control and combines two of the major features observed during motor execution: trial-to-trial variability and goal-directed corrections. Optimal feedback as a theory for motor control predicts that noise and variations are only corrected within task-relevant dimensions (that can interfere with the goal of the action), whereas they can accumulate and increase within the non-relevant

dimensions (that do not interfere with the goal of the task); Todorov and Jordan called this the “minimal intervention principle” (Todorov and Jordan, 2002).

Although behavioral and neuronal variability have been classically seen as noise, a source of decreased behavioral performance and noisy neuronal computations, more recent frameworks propose that the generation and modulation of variability can be seen as part of the mechanisms of motor control (Faisal et al., 2008; Maimon and Assad, 2009; Stein et al., 2005). Noise and variability can increase during periods of exploration and are crucial for the acquisition of skills. Nevertheless, excessive noise in the sensorimotor system may also promote a level of variability that limits the acquisition of skill (Faisal et al., 2008). Smits-Engelsman and Wilson reviewed data from a developmental coordination disorder, showing that high levels of neural noise can impair the ability to develop a reliable body schema and to implicitly learn the relationship between motor output and feedback signals (Smits-Engelsman and Wilson, 2013). These studies indicate that variability within the features that are relevant for the task outcome is a critical for learning a motor skill, and that it can be differentially modulated as training progresses.

As discussed previously in Chapter 2, we observe an increased behavioral variability within the outcome-relevant dimension during the early stages of learning that is reduced throughout learning. This was not observed within the non-relevant dimensions. Concomitantly, neuronal variability is high early in training and decreases as the animal progresses through the task. These modulations of neuronal variability are exclusively correlated with the variability observed on the outcome-relevant dimensions and not on the not-relevant ones. These data are in agreement with the minimal intervention principle by Todorov and

Jordan (Todorov and Jordan, 2002), since only the dimension that affects the result of the movement (sequence frequency) is optimized, and variability is allowed to accumulate in non-relevant dimensions (sequence length/duration). The dynamics of neuronal variability also support the principle of minimal intervention, as they specifically correlate with variability of the outcome-relevant dimension. Furthermore, we demonstrate that the refinement of outcome-relevant behavior dimensions is dependent on functional plasticity within striatal circuits.

Evidence from our work and from several other studies support the view that the basal ganglia and the corticostriatal loops are critical for selection during motor learning. According to Thorndike's "law of effect", a behavior followed by pleasant events is more likely to be reinforced and repeated in the future (Thorndike, 1898). This view is foundational to operant conditioning and reinforcement learning and has been thoroughly demonstrated in behavior studies. However, it is not possible to select specific muscle activity patterns. With the recent advances in brain-machine-interface (BMI) paradigms, we are now able to volitionally control neuronal patterns and activity without resorting to behavior as an indirect readout of neuronal activity. In Chapter 3, we investigated if neuronal selection could be directly promoted, using phasic activation of dopaminergic cells as a reinforcer for a specific outcome-relevant cortical pattern. Closed-loop BMI paradigms permit a direct testing of relevant aspects of natural motor learning in highly controlled experiments (Orsborn and Carmena, 2013). From our study we conclude that the phasic activation of dopaminergic cells, previously shown to bias the behavior of animals in self stimulation operant tasks (Olds, 1958; Rossi et al., 2013; Wise, 1981), is also sufficient to promote the selection of specific neuronal patterns. Our findings support the

theory of selection during motor learning, in accordance with the neuronal group selection and neuronal darwinism theories (Changeux, 1989; Costa, 2011; Edelman, 1987).

The view presented in this dissertation on the selection of outcome-relevant features, through the cortico-basal ganglia circuits, can be expanded as a more general feature of brain function. Beyond motor memory and the learning of actions, these frameworks might also apply to other types of memories and learning paradigms. Episodic (or declarative) memory is characterized by the descriptive report of facts regarding and event, and its consolidation is selective for specific aspects of the memory, usually associated with arousal states (McGaugh, 2013). We hypothesize that non-motor memories can be split in smaller parts with different relevance levels. The most relevant parts of the memory would be associated with lower variability levels and undergo a stronger consolidation. Fractions of the memory associated with high variability would not undergo the same consolidation, remain labile and eventually be lost with time. This framework might help promote a stronger characterization of the requirements of each task independently of the type of memory tested, using abstract concepts to provide a systematic description of the controllable dimensions, and potentially allow for a better understanding of brain function.

REFERENCES

Aldridge, J.W., Gilman, S., and Dauth, G. (1990). Spontaneous neuronal unit activity in the primate basal ganglia and the effects of precentral cerebral cortical ablations. Brain Res. *516*, 46–56.

Arsalidou, M., Duerden, E.G., and Taylor, M.J. (2013). The centre of the brain: topographical model of motor, cognitive, affective, and somatosensory functions of the basal ganglia. Hum Brain Mapp *34*, 3031–3054.

Asanuma, H., and Keller, A. (1991). Neuronal mechanisms of motor learning in mammals. Neuroreport *2*, 217–224.

Asanuma, H., and Pavlides, C. (1997). Neurobiological basis of motor learning in mammals. Neuroreport *8*, i–vi.

Asanuma, H., Fernandez, J., and Scheibel, M.E. (1974). Characteristics of projections from the nucleus ventralis lateralis to the motor cortex in the cats: an anatomical and physiological study. Experimental Brain ….

Barnes, T.D., Kubota, Y., Hu, D., Jin, D.Z., and Graybiel, A.M. (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. Nature *437*, 1158–1161.

Bolam, J.P., Hanley, J.J., Booth, P.A., and Bevan, M.D. (2000). Synaptic organisation of the basal ganglia. J. Anat. *196 (Pt 4)*, 527–542.

Bottjer, S.W., Miesner, E.A., and Arnold, A.P. (1984). Forebrain lesions disrupt development but not maintenance of song in passerine birds. Science *224*, 901–903.

Brainard, M.S., and Doupe, A.J. (2002). What songbirds teach us about learning. Nature *417*, 351–358.

Carlsson, A. (1972). Biochemical and pharmacological aspects of Parkinsonism. Acta Neurol. Scand., Suppl.C *51*, 11–42.

Changeux, J.-P. (1989). Neuronal models of cognitive functions. Cognition *33*, 63–109.

Costa, R.M. (2011). A selectionist account of de novo action learning. Curr. Opin. Neurobiol. *21*, 579–586.

Costa, R.M., Cohen, D., and Nicolelis, M.A.L. (2004). Differential corticostriatal plasticity during fast and slow motor skill learning in mice. Curr. Biol. *14*, 1124–1134.

Costa, R.M., Lin, S.-C., Sotnikova, T.D., Cyr, M., Gainetdinov, R.R., Caron, M.G., and Nicolelis, M.A.L. (2006). Rapid alterations in corticostriatal ensemble coordination during acute dopamine-dependent motor dysfunction. Neuron *52*, 359–369.

Doupe, A.J., and Kuhl, P.K. (1999). Birdsong and human speech: common themes and mechanisms. Annu. Rev. Neurosci. *22*, 567–631.

Dragoi, G., and Tonegawa, S. (2010). Preplay of future place cell sequences by hippocampal cellular assemblies. Nature.

Eckart, M.T., Huelse-Matia, M.C., McDonald, R.S., and Schwarting, R.K.W. (2010). 6-hydroxydopamine lesions in the rat neostriatum impair sequential learning in a serial reaction time task. Neurotox Res *17*, 287–298.

Edelman, G.M. (1987). Neural Darwinism: The theory of neuronal group selection. (Basic Books).

Elias, S., Ritov, Y., and Bergman, H. (2008). Balance of increases and decreases in firing rate of the spontaneous activity of basal ganglia high-frequency discharge neurons. J. Neurophysiol. *100*, 3086–3104.

Faisal, A.A., Selen, L.P.J., and Wolpert, D.M. (2008). Noise in the nervous system. Nat. Rev. Neurosci. *9*, 292–303.

Favorov, O., Sakamoto, T., and Asanuma, H. (1988). Functional role of corticoperipheral loop circuits during voluntary movements in the monkey: a preferential bias theory. J. Neurosci. *8*, 3266–3277.

Fee, M.S., and Scharff, C. (2010). The songbird as a model for the generation and learning of complex sequential behaviors. Ilar J *51*, 362–377.

Fetz, E.E., Perlmutter, S.I., Prut, Y., Seki, K., and Votaw, S. (2002). Roles of primate spinal interneurons in preparation and execution of voluntary hand movement. Brain Res. Brain Res. Rev. *40*, 53–65.

Flaherty, A.W., and Graybiel, A.M. (1991). Corticostriatal transformations in the primate somatosensory system. Projections from physiologically mapped body-part representations. J. Neurophysiol. *66*, 1249–1263.

Flaherty, A.W., and Graybiel, A.M. (1993). Two input systems for body representations in the primate striatal matrix: experimental evidence in the squirrel monkey. J. Neurosci. *13*, 1120–1137.

Friel, K.M., and Nudo, R.J. (1998). Recovery of motor function after focal cortical injury in primates: compensatory movement patterns used during rehabilitative training. Somatosens Mot Res *15*, 173–189.

Georgopoulos, A.P. (1996). On the translation of directional motor cortical commands to activation of muscles via spinal interneuronal systems. Brain Res Cogn Brain Res *3*, 151–155.

Gerbner, M. (1972). Frontal lobe and motivation of learned behavior. Acta Neurobiol Exp (Wars) *32*, 673–687.

Gerfen, C.R., and Surmeier, D.J. (2011). Modulation of striatal projection systems by dopamine. Annu. Rev. Neurosci. *34*, 441–466.

Graybiel, A.M. (1998). The basal ganglia and chunking of action repertoires. Neurobiol Learn Mem *70*, 119–136.

Grillner, S., Robertson, B., and Stephenson-Jones, M. (2013). The evolutionary origin of the vertebrate basal ganglia and its role in action selection. J Physiol (Lond) *591*, 5425–5431.

Haber, S.N. (2003). The primate basal ganglia: parallel and integrative networks. Journal of Chemical Neuroanatomy *26*, 317–330.

Haber, S.N. (2014). The place of dopamine in the cortico-basal ganglia circuit. Neurosci. *282C*, 248–257.

Hikosaka, O. (1998). Neural systems for control of voluntary action--a hypothesis. Adv. Biophys. *35*, 81–102.

Hikosaka, O., Miyashita, K., Miyachi, S., Sakai, K., and Lu, X. (1998). Differential roles of the frontal cortex, basal ganglia, and cerebellum in visuomotor sequence learning. Neurobiol Learn Mem *70*, 137–149.

Hikosaka, O., Rand, M.K., Miyachi, S., and Miyashita, K. (1995). Learning of sequential movements in the monkey: process of learning and retention of memory. J. Neurophysiol. *74*, 1652–1661.

Hoffman, D.S., and Strick, P.L. (1995). Effects of a primary motor cortex lesion on step-tracking movements of the wrist. J. Neurophysiol. *73*, 891–895.

Hunnicutt, B.J., Long, B.R., Kusefoglu, D., Gertz, K.J., Zhong, H., and Mao, T. (2014). A comprehensive thalamocortical projection map at the mesoscopic level. Nat. Neurosci. *17*, 1276–1285.

Jáidar, O., Carrillo-Reid, L., Hernández, A., Drucker-Colin, R., Bargas, J., and Hernández-Cruz, A. (2010). Dynamics of the Parkinsonian striatal microcircuit: entrainment into a dominant network state. J. Neurosci. *30*, 11326–11336.

Jin, X., and Costa, R.M. (2010). Start/stop signals emerge in nigrostriatal circuits during sequence learning. Nature *466*, 457–462.

Kao, M.H., Doupe, A.J., and Brainard, M.S. (2005). Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. Nature *433*, 638–643.

Karni, A., Meyer, G., Rey-Hipolito, C., Jezzard, P., Adams, M.M., Turner, R., and Ungerleider, L.G. (1998). The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex. Proc. Natl. Acad. Sci. USA *95*, 861–868.

Kitago, T., and Krakauer, J.W. (2013). Motor learning principles for neurorehabilitation. Handb Clin Neurol *110*, 93–103.

Kornhuber, H.H. (1978). Cortex, basal ganglia and cerebellum in motor control. Electroencephalogr Clin Neurophysiol Suppl 449–455.

Kotz, S.A., Schwartze, M., and Schmidt-Kassow, M. (2009). Non-motor basal ganglia functions: a review and proposal for a model of sensory predictability in auditory language perception. Cortex *45*, 982–990.

Kropotov, J.D., and Etlinger, S.C. (1999). Selection of actions in the basal ganglia-thalamocortical circuits: review and model. Int J Psychophysiol *31*, 197–217.

Luft, A.R., and Buitrago, M.M. (2005). Stages of motor skill learning. Mol Neurobiol *32*, 205–216.

Maimon, G., and Assad, J.A. (2009). Beyond Poisson: increased spike-time regularity across primate parietal cortex. Neuron *62*, 426–440.

Mandelblat-Cerf, Y., Paz, R., and Vaadia, E. (2009). Trial-to-trial variability of single cells in motor cortices is dynamically modified during visuomotor adaptation. J. Neurosci. *29*, 15053–15062.

Marchand, W.R. (2010). Cortico-basal ganglia circuitry: a review of key research and implications for functional connectivity studies of mood and anxiety disorders. Brain Struct Funct *215*, 73–96.

Marler, P. (1970). Birdsong and speech development: could there be parallels? Am. Sci. *58*, 669–673.

Marler, P. (1981). Birdsong: The acquisition of a learned motor skill. Trends Neurosci *4*, 88–94.

Marsden, C.D. (1984). Which motor disorder in Parkinson's disease indicates the true motor function of the basal ganglia? Ciba Found. Symp. *107*, 225–241.

McGaugh, J.L. (2013). Making lasting memories: remembering the significant. Proc. Natl. Acad. Sci. USA *110 Suppl 2*, 10402–10407.

McGeorge, A.J., and Faull, R.L. (1989). The organization of the projection from the cerebral cortex to the striatum in the rat. Neurosci. *29*, 503–537.

Miyachi, S., Hikosaka, O., Miyashita, K., Kárádi, Z., and Rand, M.K. (1997). Differential roles of monkey striatum in learning of sequential hand movement. Exp. Brain Res. *115*, 1–5.

Mogenson, G.J., Jones, D.L., and Yim, C.Y. (1980). From motivation to action: functional interface between the limbic system and the motor system. Prog Neurobiol *14*, 69–97.

Moussa, R., Poucet, B., Amalric, M., and Sargolini, F. (2011). Contributions of dorsal striatal subregions to spatial alternation behavior. Learn Mem *18*, 444–451.

Olds, J. (1958). Self-Stimulation of the Brain Its Use To Study Local Effects of Hunger, Sex, and Drugs. Science *127*, 315–324.

Orsborn, A.L., and Carmena, J.M. (2013). Creating new functional circuits for action via brain-machine interfaces. Front Comput Neurosci *7*, 157.

Pavlides, C., Miyashita, E., and Asanuma, H. (1993). Projection from the sensory to the motor cortex is important in learning motor skills in the monkey. J. Neurophysiol. *70*, 733–741.

Plenz, D., and Kital, S.T. (1999). A basal ganglia pacemaker formed by the subthalamic nucleus and external globus pallidus. Nature *400*, 677–682.

Porter, L.L., and Sakamoto, K. (1988). Organization and synaptic relationships of the projection from the primary sensory to the primary motor cortex in the cat. J. Comp. Neurol. *271*, 387–396.

Rossi, M.A., Sukharnikova, T., Hayrapetyan, V.Y., Yang, L., and Yin, H.H. (2013). Operant Self-Stimulation of Dopamine Neurons in the Substantia Nigra. PLoS ONE *8*, e65799.

Sakata, J.T., and Brainard, M.S. (2006). Real-time contributions of auditory feedback to avian vocal motor control. J. Neurosci. *26*, 9619–9628.

Scharff, C., and Nottebohm, F. (1991). A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. J. Neurosci. *11*, 2896–2913.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. Science *275*, 1593–1599.

Seo, M., Lee, E., and Averbeck, B.B. (2012). Action Selection and Action Value in Frontal-Striatal Circuits. Neuron *74*, 947–960.

Shires, J., Joshi, S., and Basso, M.A. (2010). Shedding new light on the role of the basal ganglia-

superior colliculus pathway in eye movements. Curr. Opin. Neurobiol. *20*, 717–725.

Smits-Engelsman, B.C.M., and Wilson, P.H. (2013). Noise, variability, and motor performance in developmental coordination disorder. Dev Med Child Neurol *55 Suppl 4*, 69–72.

Sohrabji, F., Nordeen, E.J., and Nordeen, K.W. (1990). Selective impairment of song learning following lesions of a forebrain nucleus in the juvenile zebra finch. Behav. Neural Biol. *53*, 51–63.

Sporns, O., and Tononi, G. (1994). Selectionism and the Brain (Academic Press).

Stein, R.B., Gossen, E.R., and Jones, K.E. (2005). Neuronal variability: noise or part of the signal? Nat. Rev. Neurosci. *6*, 389–397.

Surmeier, D.J., Carrillo-Reid, L., and Bargas, J. (2011). Dopaminergic modulation of striatal neurons, circuits, and assemblies. Neurosci. *198*, 3–18.

Sutton, R.S., and Barto, A.G. (1998). Reinforcement Learning (MIT Press).

Thorndike, E.L. (1898). Animal intelligence: An experimental study of the associative processes in animals. The Psychological Review: Monograph Supplements *2*, i–109.

Thorpe, W.H. (1958). The learning of song patterns by birds, with especial reference to the song of the chaffinch Fringilla coelebs. Ibis *100*, 535–570.

Todorov, E., and Jordan, M.I. (2002). Optimal feedback control as a theory of motor coordination. Nat. Neurosci. *5*, 1226–1235.

Tumer, E.C., and Brainard, M.S. (2007). Performance variability enables adaptive plasticity of “crystallized” adult birdsong. Nature *450*, 1240–1244.

Ungerleider, L.G., Doyon, J., and Karni, A. (2002). Imaging brain plasticity during motor skill learning. Neurobiol Learn Mem *78*, 553–564.

Voorn, P., Vanderschuren, L.J.M.J., Groenewegen, H.J., Robbins, T.W., and Pennartz, C.M.A. (2004). Putting a spin on the dorsal-ventral divide of the striatum. Trends Neurosci *27*, 468–474.

Whishaw, I.Q., Mittleman, G., Bunch, S.T., and Dunnett, S.B. (1987). Impairments in the acquisition, retention and selection of spatial navigation strategies after medial caudate-putamen lesions in rats. Behavioural Brain Research *24*, 125–138.

Williams, H. (2004). Birdsong and singing behavior. Ann. N. Y. Acad. Sci. *1016*, 1–30.

Wise, R.A. (1981). Intracranial self-stimulation: mapping against the lateral boundaries of the dopaminergic cells of the substantia nigra. Brain Res. *213*, 190–194.

Wu, H.G., Miyamoto, Y.R., Gonzalez Castro, L.N., Olveczky, B.P., and Smith, M.A. (2014). Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. Nat. Neurosci. *17*, 312–321.

Yin, H.H., Mulcare, S.P., Hilário, M.R.F., Clouse, E., Holloway, T., Davis, M.I., Hansson, A.C., Lovinger, D.M., and Costa, R.M. (2009). Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. Nat. Neurosci. *12*, 333–341.

ITQB-UNL | Av. da República, 2780-157 Oeiras, Portugal
Tel (+351) 214 469 100 | Fax (+351) 214 411 277

www.itqb.unl.pt